NANO- AND MICROELECTROMECHANICAL SYSTEMS

Fundamentals of Nano- and Microengineering

VISIT…

A book in the
Nano- and Microscience, Engineering,
Technology and Medicine Series

NANO- AND MICROELECTROMECHANICAL SYSTEMS

Fundamentals of Nano- and Microengineering

Sergey Edward Lyshevski

CRC Press
Boca Raton London New York Washington, D.C.

Library of Congress Cataloging-in-Publication Data

Lyshevski, Sergey Edward.
Nano- and microelectromechanical systems : fundamentals of nano- and microengineering / Sergey Edward Lyshevski.
p. cm. -- (Nano- and microscience, engineering, technology, and medicine series)
Includes index.
Includes bibliographical references and index.
ISBN 0-8493-916-6 (alk. paper)
1. Microelectromechanical systems. 1. Title. II. Series.

TK7875 .L96 2000
621.381—dc201 00-057953
CIP

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

Neither this book nor any part may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, microfilming, and recording, or by any information storage or retrieval system, without prior permission in writing from the publisher.

The consent of CRC Press LLC does not extend to copying for general distribution, for promotion, for creating new works, or for resale. Specific permission must be obtained in writing from CRC Press LLC for such copying.

Direct all inquiries to CRC Press LLC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation, without intent to infringe.

© 2001 by CRC Press LLC

No claim to original U.S. Government works
International Standard Book Number 0-8493-916-6
Library of Congress Card Number 00-057953
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

To my family

PREFACE

This book is designed for a one-semester course on Nano- and Microelectromechanical Systems or Nano- and Microengineering. A typical background needed includes calculus, electromagnetics, and physics. The purpose of this book is to bring together in one place the various methods, techniques, and technologies that students and engineers need in solving a wide array of engineering problems in formulation, modeling, analysis, design, and optimization of high-performance microelectromechanical and nanoelectromechanical systems (MEMS and NEMS). This book is not intended to cover fabrication aspects and technologies because a great number of books are available. At the same time, extremely important issues in analysis, design, modeling, optimization, and simulation of NEMS and MEMS have not been comprehensively covered in the existing literature.

Twenty first century nano- and microtechnology revolution will lead to fundamental breakthroughs in the way materials, devices, and systems are understood, designed, function, manufactured, and used. Nanoengineering and nanotechnology will change the nature of the majority of the human-made structures, devices, and systems. Current technological needs and trends include technology development and transfer, manufacturing and deployment, implementation and testing, modeling and characterization, design and optimization, simulation and analysis of complex nano- and microscale devices (for example, molecular computers, logic gates and switches, actuators and sensors, digital and analog integrated circuits, et cetera). Current developments have been focused on analysis and synthesis of molecular structures and devices which will lead to revolutionary breakthroughs in the data processing and computing, data storage and imaging, quantum computing and molecular intelligent automata, etc. Micro- and nanoengineering and science lead to fundamental breakthroughs in the way materials, devices and systems are understood, designed, function, manufactured, and used. High-performance MEMS and NEMS, micro- and nanoscale structures and devices will be widely used in nanocomputers, medicine (nanosurgery and nanotherapy, nonrejectable artificial organ design and implants, drug delivery and diagnosis), biotechnology (genome synthesis), etc.

New phenomena in nano- and microelectromechanics, physics and chemistry, benchmarking nanomanufacturing and control of complex molecular structures, design of large-scale architectures and optimization, among other problems must be addressed and studied. The major objective of this book is the development of basic theory (through multidisciplinary fundamental and applied research) to achieve full understanding, optimize, and control properties and behavior of a wide range of NEMS and MEMS. This will lead to new advances and will allow the designer to comprehensively solve a number of long-standing problems in analysis and

control, modeling and simulation, structural optimization and virtual prototyping, packaging and fabrication, as well as implementation and deployment of novel NEMS and MEMS. In addition to technological developments and manufacturing (fabrication), the ability to synthesize and optimize NEMS and MEMS depends on the analytical and numerical methods, and the current concepts and conventional technologies cannot be straightforwardly applied due to the highest degree of complexity as well as novel phenomena. Current activities have been centered in development and application of a variety of experimental techniques trying to attain the characterization of mechanical (structural and thermal), electromagnetic (conductivity and susceptibility, permittivity and permeability, charge and current densities, propagation and radiation), optical, and other properties of NEMS and MEMS. It has been found that CMOS, surface micromachining and photolithography, near-field optical microscopy and magneto-optics, as well as other leading-edge technologies and processes to some extent can be applied and adapted to manufacture nano- and microscale structures and devices. However, advanced interdisciplinary research must be carried out to design, develop, and implement high-performance NEMS and MEMS. Our objectives are to expand the frontiers of the NEMS- and MEMS-based research through pioneering fundamental and applied multidisciplinary studies and developments. Rather than designing nano- and microscale components (integrated circuits and antennas, electromechanical and opto-electromechanical actuators and sensors), the emphasis will be given to the synthesis of the integrated large-scale systems. It must be emphasized that the author feels quite strongly that the individual nano- and microscale structures must be synthesized, thoroughly analyzed, and studied. We will consider NEMS and MEMS as the large-scale highly coupled systems, and the synthesis of groups of cooperative multi-agent NEMS and MEMS can be achieved using hierarchical structural and algorithmic optimization methods. The optimality of NEMS and MEMS should be guaranteed with respect to a certain performance objectives (manufacturing and packaging, cost and maintenance, size and weight, efficiency and performance, affordability and reliability, survivability and integrity, et cetera).

Nanoengineering is a very challenging field due to the complex multidisciplinary nature (engineering and physics, biology and chemistry, technology and material science, mathematics and medicine). This book introduces the focused fundamentals of nanoelectromechanics to initiate and stress, accelerate and perform the basic and applied research in NEMS and MEMS. Many large-scale systems are too complex to be studied and optimized analytically, and usually the available information is not sufficient to derive and obtain performance functionals. Therefore, the stochastic gradient descent and nonparametric methods can be applied using the decision variables with conflicting specifications and requirements imposed. In many applications there is a need to design high-performance intelligent NEMS and MEMS to accomplish the following functions:

- programming and self-testing;
- collection, compiling, and processing information (sensing – data accumulation (storage) – processing);
- multivariable embedded high-density array coordinated control;
- calculation and decision making with outcomes prediction;
- actuation and control.

The fundamental goal of this book is to develop the basic theoretical foundations in order to design and develop, analyze and prototype high-performance NEMS and MEMS. This book is focused on the development of fundamental theory of NEMS and MEMS, as well as their components and structures, using advanced multidisciplinary basic and applied developments. In particular, it will be illustrated how to perform the comprehensive studies with analysis of the processes, phenomena, and relevant properties at nano- and micro-scales, development of NEMS and MEMS architectures, physical representations, structural design and optimization, etc. It is the author's goal to substantially contribute to these basic issues, and the integration of these problems in the context of specific applications will be addressed. The primary emphasis will be on the development of basic theory to attain fundamental understanding of NEMS and MEMS, processes in nano- and micro-scale structures, as well as the application of the developed theory. Using the molecular technology, one can design and manufacture the atomic-scale devices with atomic precision using the atomic building blocks, design nano-scale devices ranging from electromechanical motion devices (translational and rotational actuators and sensors, logic and switches, registers) to nano-scale integrated circuits (diodes and transistors, logic gates and switches, resistors and inductors, capacitors). These devices will be widely used in medicine and avionics, transportation and power, and many other areas.

The leading-edge research in nanosystems is focused on different technologies and processes. As an example, the discovery of carbon-based nanoelectronics (carbon nanotubes are made from individual molecules) is the revolutionary breakthrough in nanoelectronics and nanocomputers, information technology and medicine, health and national security. In particular, fibers made using carbon nanotubes (molecular wires) more than 100 times stronger than steel and weighing 5 times less, have conductivity 5 times greater than silver, and transmit heat better than diamond. Carbon nanotubes are used as the molecular wires. Furthermore, using carbon molecules, first single molecule transistors were built. It should be emphasized that the current technology allows one to fill carbon nanotubes with other media (metals, organic and inorganic materials, et cetera).

The research in nano- and microtechnologies will lead to breakthroughs in information technology and manufacturing, medicine and health, environment and energy, avionics and transportation, national security and other areas of the greatest national importance. Through interdisciplinary synergism, this book is focused on fundamental studies of phenomena and

processes in NEMS and MEMS, synthesis of nano- and micro-scale devices and systems, design of building blocks and components (which will lead to efficient and affordable manufacturing of high-performance NEMS and MEMS), study of molecular structures and their control, NEMS and MEMS architectures, etc. We will discuss the application and impact of nano- and micro-scale structures, devices, and systems to information technology, nanobiotechnology and medicine, nanomanufacturing and environment, power and energy systems, health and national security, avionics and transportation.

Acknowledgments

Many people contributed to this book. First thanks go to my beloved family. I would like to express my sincere acknowledgments and gratitude to many colleagues and students. It gives me great pleasure to acknowledge the help I received from many people in the preparation of this book. The outstanding team of the CRC Press, especially Nora Konopka (Acquisition Editor Electrical Engineering) and William Heyward (Project Editor), tremendously helped and assisted me providing valuable and deeply treasured feedback. Many thanks for all of you.

CONTENTS

1. Nano- and Microengineering, and Nano- and Microtechnologies

1.1. Introduction
1.2. Biological Analogies
1.3. Nano- and Microelectromechanical Systems
1.4. Applications of Nano- and Microelectromechanical Systems
1.5. Nano- and Microelectromechanical Systems
1.6. Introduction to MEMS Fabrication, Assembling, and Packaging

2. Mathematical Models and Design of Nano- and Microelectromechanical Systems

2.1. Nano- and Microelectromechanical Systems Architecture
2.2. Electromagnetics and its Application For Nano- and Microscale Electromechanical Motion Devices
2.3. Classical Mechanics and its Application
2.3.1. Newtonian Mechanics
2.3.2. Lagrange Equations of Motion
2.3.3. Hamilton Equations of Motion
2.4. Atomic Structures and Quantum Mechanics
2.5. Molecular and Nanostructure Dynamics
2.5.1. Schrödinger Equation and Wavefunction Theory
2.5.2. Density Functional Theory
2.5.3. Nanostructures and Molecular Dynamics
2.6. Molecular Wires and Molecular Circuits
2.7 Thermoanalysis and Heat Equation

3. Structural Design, Modeling, and Simulation

3.1. Nano- and Microelectromechanical Systems
3.1.1. Carbon Nanotubes and Nanodevices
3.1.2. Microelectromechanical Systems and Microdevices
3.2. Structural Synthesis of Nano- and Microelectromechanical Actuators and Sensors
3.2.1. Configurations and Structural Synthesis of Motion Nano- and Microstructures (actuators and Sensors)
3.2.2. Algebra of Sets
3.3. Direct-Current Micromachines
3.4. Induction Motors
3.4.1. Two-Phase Induction Motors
3.4.2. Three-Phase Induction Motors
3.5. Microscale Synchronous Machines
3.5.1. Single-Phase Reluctance Motors
3.5.2. Permanent-Magnet Synchronous Machines

3.6. Microscale Permanent-Magnet Stepper Motors
3.6.1. Mathematical Model in the Machine Variables
3.6.2. Mathematical Models of Permanent-Magnet Stepper Motors in the Rotor and Synchronous Reference Frames
3.7. Nanomachines: Nanomotors and Nanogenerators

4. Control of Nano- and Microelectromechanical Systems
4.1. Fundamentals of Electromagnetic Radiation and Antennas in Nano- and Microscale Electromechanical Systems
4.2. Design of Closed-Loop Nano- and Microelectromechanical Systems Using the Lyapunov Stability Theory

4.3. Introduction to Intelligent Control of Nano- and Microelectromechanical Systems

CHAPTER 1

NANO- AND MICROENGINEERING, AND NANO- AND MICROTECHNOLOGIES

1.1. INTRODUCTION

The development and deployment of NEMS and MEMS are critical to the U.S. economy and society because nano- and microtechnologies will lead to major breakthroughs in information technology and computers, medicine and health, manufacturing and transportation, power and energy systems, and avionics and national security. NEMS and MEMS have important impacts in medicine and bioengineering (DNA and genetic code analysis and synthesis, drug delivery, diagnostics, and imaging), bio and information technologies, avionics, and aerospace (nano- and microscale actuators and sensors, smart reconfigurable geometry wings and blades, space-based flexible structures, and microgyroscopes), automotive systems and transportation (sensors and actuators, accelerometers), manufacturing and fabrication, public safety, etc. During the last years, the government and the high-technology industry have heavily funded basic and applied research in NEMS and MEMS due to the current and potential rapidly growing positive direct and indirect social and economic impacts.

Nano- and microengineering are the fundamental theory, engineering practice, and leading-edge technologies in analysis, design, optimization, and fabrication of NEMS and MEMS, nano- and microscale structures, devices, and subsystems. The studied nano- and microscale structures and devices have dimensions of nano- and micrometers.

To support the nano- and microtechnologies, basic and applied research and development must be performed. Nanoengineering studies nano- and microscale-size materials and structures, as well as devices and systems, whose structures and components exhibit novel physical (electromagnetic and electromechanical), chemical, and biological properties, phenomena, and processes. The dimensions of nanosystems and their components are 10^{-10} m (molecule size) to 10^{-7} m; that is, 0.1 to 100 nanometers. Studying nanostructures, one concentrates one's attention on the atomic and molecular levels, manufacturing and fabrication, control and dynamics, augmentation and structural integration, application and large-scale system synthesis, et cetera. Reducing the dimensions of systems leads to the application of novel materials (carbon nanotubes, quantum wires and dots). The problems to be solved range from mass-production and assembling (fabrication) of nanostructures at the atomic/molecular scale (e.g., nanostructured electronics and actuators/sensors) with the desired properties. It is essential to design novel nanodevices such as nanotransistors and nanodiodes, nanoswitches and nanologic gates, in order to design nanoscale computers with terascale capabilities. All living biological

systems function due to molecular interactions of different subsystems. The molecular building blocks (proteins and nucleic acids, lipids and carbohydrates, DNA and RNA) can be viewed as inspiring possible strategy on how to design high-performance NEMS and MEMS that possess the properties and characteristics needed. Analytical and numerical methods are available to analyze the dynamics and three-dimensional geometry, bonding, and other features of atoms and molecules. Thus, electromagnetic and mechanical, as well as other physical and chemical properties can be studied.

Nanostructures and nanosystems will be widely used in medicine and health. Among possible applications of nanotechnology are: drug synthesis and drug delivery (the therapeutic potential will be enormously enhanced due to direct effective delivery of new types of drugs to the specified body sites), nanosurgery and nanotherapy, genome synthesis and diagnostics, nanoscale actuators and sensors (disease diagnosis and prevention), nonrejectable nano-artificial organs design and implant, and design of high-performance nanomaterials.

It is obvious that nano- and microtechnologies drastically change the fabrication and manufacturing of materials, devices, and systems through:

- predictable properties of nano composites and materials (e.g., light weight and high strength, thermal stability, low volume and size, extremely high power, torque, force, charge and current densities, specified thermal conductivity and resistivity, et cetera),
- virtual prototyping (design cycle, cost, and maintenance reduction),
- improved accuracy and precision, reliability and durability,
- higher degree of efficiency and capability, flexibility and integrity, supportability and affordability, survivability and redundancy,
- improved stability and robustness,
- higher degree of safety,
- environmental competitiveness.

Foreseen by Richard Feyman, the term “nanotechnology” was first used by N. Taniguchi in his 1974 paper, "On the basic concept of nanotechnology." In the last two decades, nanoengineering and nanomanufacturing have been popularized by Eric Drexler through the Foresight Institute.

Advancing miniaturization towards the molecular level with the ultimate goal to design and manufacture nanocomputers and nanomanipulators (nanoassemblers), large-scale intelligent NEMS and MEMS (which have nanocomputers as the core components), the designer faces a great number of unsolved problems.

Possible basic concepts in the development of nanocomputers are listed below. Mechanical “computers” have the richest history traced thousand years back. While the most creative theories and machines have been developed and demonstrated, the feasibility of mechanical nanocomputers is questioned by some researchers due to the number of mechanical components (which are needed to be controlled), as well as due to unsolved

manufacturing (assembling) and technological difficulties. Chemical nanocomputers can be designed based upon the processing information by making or breaking chemical bonds, and storing the information in the resulting chemical. In contrast, in quantum nanocomputers, the information can be represented by a quantum state (e.g., the spin of the atom can be controlled by the electromagnetic field).

Electronic nanocomputers can be designed using conventional concepts tested and used for the last thirty years. In particular, molecular transistors or quantum dots can be used as the basic elements. The nanoswitches (memoryless processing elements), logic gates, and registers must be manufactured on the scale of a single molecule. The so-called quantum dots are metal boxes that hold the discrete number of electrons which is changed applying the electromagnetic field. The quantum dots are arranged in the quantum dot cells. Consider the quantum dot cells which have five dots and two quantum dots with electrons. Two different states are illustrated in Figure 1.1.1 (the dashed dots contain the electron, while the white dots do not contain the electron). It is obvious that the quantum dots can be used to synthesize the logic devices.

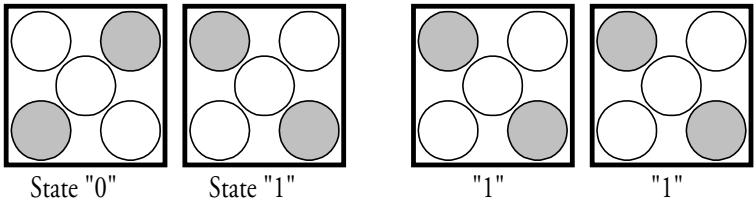


Figure 1.1.1. Quantum dots with states “0” and “1”, and “1 1” configuration

It was emphasized that as conventional electromechanical systems, nanoelectromechanical systems (actuators and other molecular devices) are controlled by changing the electromagnetic field. It becomes evident that other nanoscale structures and devices (nanodiodes and nanotransistors) are also controlled by applying the electromagnetic field (recall that the voltage and current result due to the electromagnetic field).

1.2. BIOLOGICAL ANALOGIES

Coordinated behavior and motion, visualization and sensing, motoring and decision making, memory and learning of living organisms are the results of the electrical (electromagnetic) transmission of information by neurons. One cubic centimeter of the brain contains millions of nerve cells, and these cells communicate with thousands of neurons creating data processing (communication) networks. The information from the brain to the muscles is transmitted within the milliseconds, and the baseball and football, basketball,

and tennis players calculate the speed and velocity of the ball, analyze the situation, make the decision, and respond (e.g., run or jump, throw or hit the ball, et cetera). Human central nervous system, which includes brain and spinal cord, serves as the link between the sensors (sensor receptors) and motors peripheral nervous system (effector, muscle, and gland cells). It should be emphasized that the nervous system has the following major functions: sensing, integration and decision making (computing), and motoring (actuation). Human brain consists of hindbrain (controls homeostasis and coordinate movement), midbrain (receiving, integration, and processing the sensory information), and forebrain (neural processing and integration of information, image processing, short- and long-term memories, learning functions, decision making and motor command development). The peripheral nervous system consists of the sensory system (sensory neurons transmit information from internal and external environment to the central nervous system, and motor neurons carry information from the brain or spinal cord to effectors), which supplies information from sensory receptors to the central nervous system, and the motor nervous system feeds signals (commands) from the central nervous system to muscles (effectors) and glands. The spinal cord mediates reflexes that integrate sensor inputs and motor outputs, and through the spinal cord the neurons carry information to and from the brain. The transmission of electrical signals along neurons is a very complex phenomenon. The membrane potential for a nontransmitting neuron is due to the unequal distribution of ions (sodium and potassium) across the membrane. The resting potential is maintained due to the differential ion permeability and the so-called Na^+ - K^+ pump. The stimulus changes the membrane permeability, and ion can depolarize or hyperpolarize the membrane resting potential. This potential (voltage) change is proportional to the strength of the stimulus. The stimulus is transmitted due to the axon mechanism. The nervous system is illustrated in Figure 1.2.1.

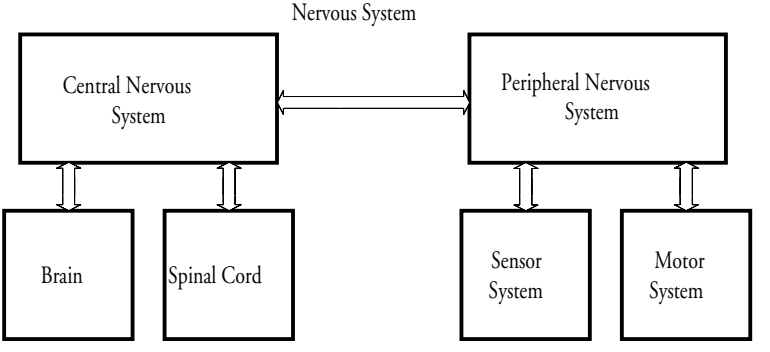


Figure 1.2.1. Vertebrate nervous system: high-level functional diagram

There is a great diversity of the nervous system organizations. The cnidarian (*hydra*) nerve net is an organized system of nerves with no central

control, and a simple nerve net can perform elementary tasks (jellyfishes swim). Echinoderms have a central nerve ring with radial nerves (for example, sea stars have central and radial nerves with nerve net). Planarians have small brains that send information through two or more nerve trunks, as illustrated in Figure 1.2.2.

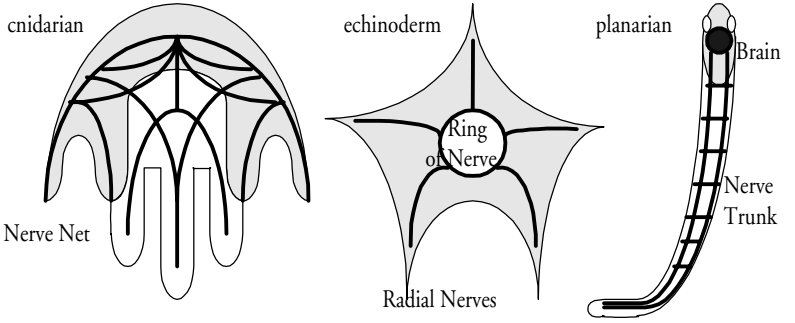


Figure 1.2.2. Overview of invertebrate nervous systems

1.3. NANO- AND MICROELECTROMECHANICAL SYSTEMS

Through biosystems analogy, a great variety of man-made electromechanical systems have been designed and made. To analyze, design, develop, and deploy novel NEMS and MEMS, the designer must synthesize advanced architectures, integrate the latest advances in nano- and microscale actuators/sensors (transducers) and smart structures, integrated circuits (ICs) and multiprocessors, materials and fabrications, structural design and optimization, modeling and simulation, et cetera. It is evident that novel optimized NEMS and MEMS architectures (with processors or multiprocessors, memory hierarchies and multiple parallelism to guarantee high-performance computing and decision making), new smart structures and actuators/sensors, ICs and antennas, as well as other subsystems play a critical role in advancing the research, developments, and implementation. In this book we discuss optimized architectures, and the research in architecture optimization will provide deep insights into how intelligent large-scale integrated NEMS and MEMS can be synthesized.

Electromechanical systems, as shown in Figure 1.3.1, can be classified as

- conventional electromechanical systems,
- microelectromechanical systems (MEMS),
- nanoelectromechanical systems (NEMS).

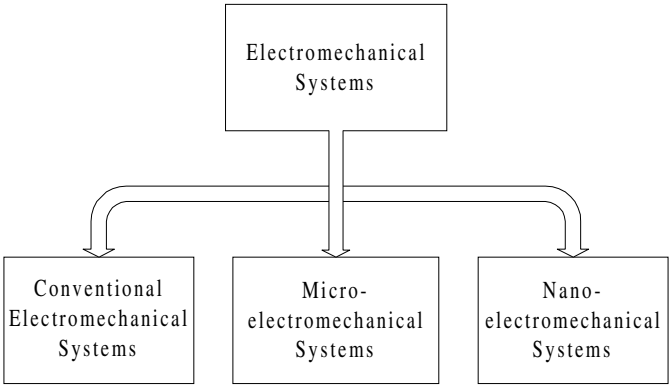


Figure 1.3.1. Classification of electromechanical systems

The operational principles and basic foundations of conventional electromechanical systems and MEMS are the same, while NEMS are studied using different concepts and theories. In fact, the designer applies the classical Lagrangian and Newtonian mechanics as well as electromagnetics (Maxwell's equations) to study conventional electromechanical systems and MEMS. In contrast, NEMS are studied using quantum theory and nanoelectromechanical concepts. Figure 1.3.2 documents the fundamental theories to study the processes and phenomena in conventional, micro, and nanoelectromechanical systems.

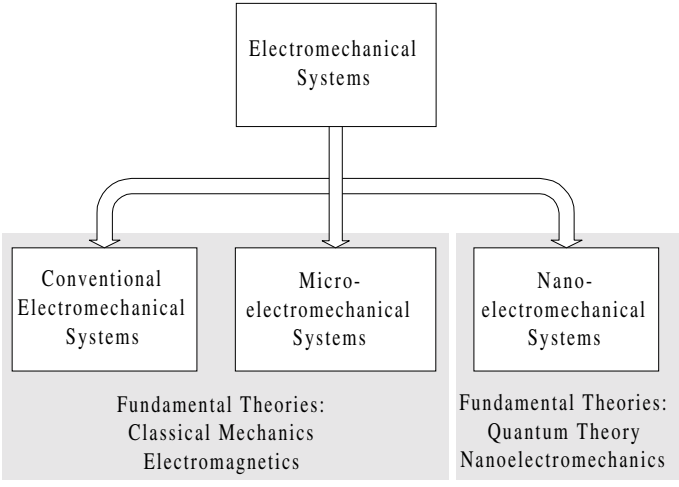


Figure 1.3.2. Fundamental theories in electromechanical systems

NEMS and MEMS integrate different structures, devices, and subsystems. The research in integration and optimization (optimized architectures and structural optimization) of these subsystems has not been instituted and performed, and end-to-end (processors – networks – input/output subsystems – ICs/antennas – actuators/sensors) performance and behavior must be studied. Through this book we will study different NEMS and MEMS architectures, and fundamental and applied theoretical concepts will be developed and documented in order to design next generation of superior high-performance NEMS and MEMS.

The large-scale NEMS and MEMS, which can integrate processor (multiprocessor) and memories, high-performance networks and input-output (IO) subsystems, are of far greater complexity than MEMS commonly used today. In particular, the large-scale NEMS and MEMS can integrate:

- thousands of nodes of high-performance actuators/sensors and smart structures controlled by ICs and antennas;
- high-performance processors or superscalar multiprocessors;
- multi-level memory and storage hierarchies with different latencies (thousands of secondary and tertiary storage devices supporting data archives);
- interconnected, distributed, heterogeneous databases;
- high-performance communication networks (robust, adaptive intelligent networks).

It must be emphasized that even the simplest nanosystems (for example, pure actuator) usually cannot function alone. For example, at least the internal or external source of energy is needed.

The complexity of large-scale NEMS and MEMS requires new fundamental and applied research and developments, and there is a critical need for coordination across a broad range of hardware and software. For example, design of advanced nano- and microscale actuators/sensors and smart structures, synthesis of optimized (balanced) architectures, development of new programming languages and compilers, performance and debugging tools, operating system and resource management, high-fidelity visualization and data representation systems, design of high-performance networks, et cetera. New algorithms and data structures, advanced system software and distributed access to very large data archives, sophisticated data mining and visualization techniques, as well as advanced data analysis are needed. In addition, advanced processor and multiprocessors are needed to achieve sustained capability required of functionally *usable* large-scale NEMS and MEMS.

The fundamental and applied research in NEMS and MEMS has been dramatically affected by the emergence of high-performance computing. Analysis and simulation of NEMS and MEMS have significant outcomes. The problems in analysis, modeling, and simulation of large-scale NEMS and MEMS that involves the complete molecular dynamics cannot be solved because the classical quantum theory cannot be feasibly applied to complex molecules or simplest nanostructures (1 nm cube of nanoactuator has thousands

of molecules). There are a number of very challenging research problems in which advanced theory and high-end computing are required to advance the theory and engineering practice. The multidisciplinary fundamentals of nanoelectromechanics must be developed to guarantee the possibility to synthesize, analyze, and fabricate high-performance NEMS and MEMS with desired (specified) performance characteristics. This will dramatically shorten the time and cost of developments of NEMS and MEMS for medical and biomedical, aerospace and automotive, electronic and manufacturing systems.

The importance of mathematical model developments and numerical analysis has been emphasized. Numerical simulation enhances, but does not substitute for fundamental research. Furthermore, meaningful and explicit simulations should be based on reliable fundamental studies and must be validated through experiments. However, it is evident that simulations lead to understanding of performance of complex NEMS and MEMS (nano- and microscale structures, devices, and sub-systems), reduce the time and cost of deriving and leveraging the NEMS and MEMS technologies from concept to device/system, and from device/system to market. Fundamental and applied research is the core of the simulation, and focused efforts must be concentrated on comprehensive modeling and advanced efficient computing.

To comprehensively study NEMS and MEMS, advanced modeling and computational tools are required primarily for 3D+ (three-dimensional geometry dynamics in time domain) data intensive modeling and simulations to study the end-to-end dynamic behavior of actuators and sensors. The mathematical models of NEMS, MEMS, and their components (structures, devices, and subsystems) must be developed. These models (augmented with efficient computational algorithms, terascale computers, and advanced software) will play the major role to simulate the design of NEMS and MEMS from virtual prototyping standpoints.

There are three broad categories of problems for which new algorithms and computational methods are critical:

1. Problems for which basic fundamental theories are developed, but the complexity of solutions is beyond the range of current and near-future computing technologies. For example, the conceptually straightforward classical quantum mechanics and molecular dynamics cannot be applied even for nanoscale actuators. In contrast, it will be illustrated that it is possible to perform robust predictive simulations of molecular-scale behavior for nano- and microscale actuators/sensors and smart structures which might contain millions of molecules.
2. Problems for which fundamental theories are not completely developed to justify direct simulations, but can be advanced or developed by advanced basic and numerical methods.
3. Problems for which the developed advanced modeling and simulation methods will produce major advances and will have a major impact. For example, 3D+ transient end-to-end behavior of NEMS and MEMS.

   For NEMS and MEMS, as well as for their devices and subsystems,

high-fidelity modeling and massive computational simulations (mathematical models designed with developed intelligent libraries and databases/archives, intelligent experimental data manipulation and storage, data grouping and correlation, visualization, data mining and interpretation) offer the promise of developing and understanding the mechanisms, phenomena and processes in order to improve efficiency and design novel high-performance NEMS and MEMS. Predictive model-based simulations require terascale computing and an unprecedented level of integration between engineering and science. These modeling and simulations will lead to new fundamental results. To model and simulate NEMS and MEMS, we augment modern quantum mechanics, electromagnetics, and electromechanics at the nano- and microscale. In particular, our goal is to develop the nanoelectromechanical theory.

One can perform the steady-state and dynamic analysis. While steady-state analysis is important, and the structural optimization to comprehend the actuators/sensors, smart structures, and antennas design can be performed, NEMS and MEMS must be analyzed in the time domain. The long-standing goal of nanoelectromechanics is to develop the basic fundamental conceptual theory in order to determine and study the interactions between actuation and sensing, computing and communication, signal processing and hierarchical data storage (memories), and other processes and phenomena in NEMS and MEMS. Using the concept of strong electromagnetic-electromechanical interactions, the fundamental nanoelectromechanical theory will be developed and applied to nanostructures and nanodevices, NEMS and MEMS to predict the performance through analytical solutions and numerical simulations. Dynamic macromodels of nodes can be developed, and single and groups of molecules can be studied. It is critical to perform this research in order to determine a number of the parameters to make accurate performance evaluation and to analyze the phenomena performing simulations and comparing experimental, modeling and simulation results.

Current advances and developments in modeling and simulation of complex phenomena in NEMS and MEMS are increasingly dependent upon new approaches to robustly map, compute, visualize, and validate the results clarifying, correlating, defining, and describing the limits between the numerical results and the qualitative-quantitative analytic analysis in order to comprehend, understand, and grasp the basic features. Simulations of NEMS and MEMS require terascale computing that will be available within a couple of years. The computational limitations and inability to develop explicit mathematical models (some nonlinear phenomena cannot be comprehended, fitted, and precisely mapped) focus advanced studies on the basic research in robust modeling and simulation under uncertainties. Robust modeling, simulation, and design are critical to advance and foster the theoretical and engineering enterprises. We focus our research on the development of the nanoelectromechanical theory in order to model and simulate large-scale NEMS and MEMS. At the subsystem level, for example, nano- and microscale actuators and sensors will be modeled and analyzed in 3D+ (three-dimensional

geometry dynamics in time domain) applying advanced numerical robust methods and algorithms. Rigorous methods for quantifying uncertainties for robust analysis should be developed. Uncertainties result due to the fact that it is impossible to explicitly comprehend the complex interacted subsystems and processes in NEMS and MEMS (actuators/sensors and smart structures, antennas, digital and analog ICs, data movement, storage and management across multilevel memory hierarchies, archives, networks and periphery), structural and environmental changes, unmeasured and unmodeled phenomena, et cetera.

To design NEMS and MEMS, we will develop analytical mathematical models. There are a number of areas where the advances must be made in order to realize the promises and benefits of modern theoretical developments recently made. For example, to perform 3D+ modeling and data intensive simulations of actuators/sensors and smart structures, we will use advanced analytical and numerical methods and algorithms (novel methods and algorithms in geometry and mesh generation, data assimilation, and dynamic adaptive mesh refinement) as well as the computationally efficient and robust MATLAB environment. There are fundamental and computational problems that have not been addressed, formulated and solved due to the complexity of large-scale NEMS and MEMS (e.g., large-scale hybrid models, limited ability to generate and visualize the massive amount of data, et cetera). Other problems include nonlinearities and uncertainties which imply fundamental limits to formulate, set up, and solve analysis and design problems. Therefore, one should develop rigorous methods and algorithms for quantifying and modeling uncertainties, 3D+ geometry and mesh generation techniques, as well as methods for adaptive robust modeling and simulations under uncertainties. A broad class of fundamental and applied problems ranging from fundamental theories (quantum mechanics and electromagnetics, electromechanics and thermodynamics, structural synthesis and optimization, optimized architecture design and control, modeling and analysis, et cetera) and numerical computing (to enable the major progress in design and virtual prototyping through the large scale simulations, data intensive computing, and visualization) will be addressed and thoroughly studied in this book. Due to the obvious limitations and the scope of this book, a great number of problems and phenomena will not be addressed and discussed (among them, fabrication and manufacturing, chemistry and material science).

1.4. APPLICATIONS OF NANO- AND MICROELECTROMECHANICAL SYSTEMS

Depending upon the specifications and requirements, objectives and applications, NEMS and MEMS must be designed. Usually, NEMS are faster and simpler, more efficient and reliable, survivable and robust compared with MEMS. However, due to the limited size and functional capabilities, one might not attain the desired characteristics. For example, consider nano-

and microscale actuators. The actuator size is determined by the force or torque densities. That is, the size is determined by the force or torque requirements and materials used. As one uses NEMS or MEMS as the logic devices, the output electric signal (voltage or current) or electromagnetic field (intensity or density) must have the specified value.

Although NEMS and MEMS have the common features, the differences must be emphasized as well. Currently, the research and developments in NEMS and molecular nanotechnology are primarily concentrated on design, modeling, simulation, and fabrication of molecular-scale devices. In contrast, MEMS are usually fabricated using other technologies, for example, complementary metal oxide semiconductor (CMOS) and lithography. The direct chip attaching technology was developed and widely deployed. Flip-chip assembly replaces wire banding to connect ICs with micro- and nanoscale actuators and sensors. The use of flip-chip technology allows one to eliminate parasitic resistance, capacitance, and inductance. This results in improvements of performance characteristics. In addition, flip-chip assembly offers advantages in the implementation of advanced flexible packaging, improving reliability and survivability, reduces weight and size, et cetera. The flip-chip assembly involves attaching actuators and sensors directly to ICs. The actuators and sensors are mounted face down with bumps on the pads that form electrical and mechanical joints to the ICs substrate. The under-fill encapsulate is then added between the chip surface and the flex circuit to achieve the high reliability demanded. Figure 1.4.1 illustrates flip-chip MEMS.

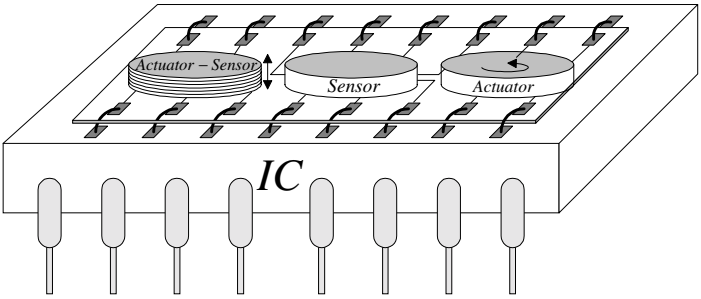


Figure 1.4.1. Flip-chip monolithic MEMS with actuators and sensors

The large-scale integrated MEMS (a single chip that can be mass-produced using the complementary metal oxide semiconductor (CMOS), photolithography, and other technologies at low cost) integrates:

- N nodes of actuators/sensors, smart structures,
- ICs and antennas,
- processor and memories,
- interconnection networks (communication busses),
- input-output (IO) systems.

Different architectures can be synthesized, and this problem is discussed

and covered in Chapter 2. One uses NEMS and MEMS to control complex systems, processes, and phenomena. A high-level functional block diagram of large-scale MEMS is illustrated in Figure 1.4.2.

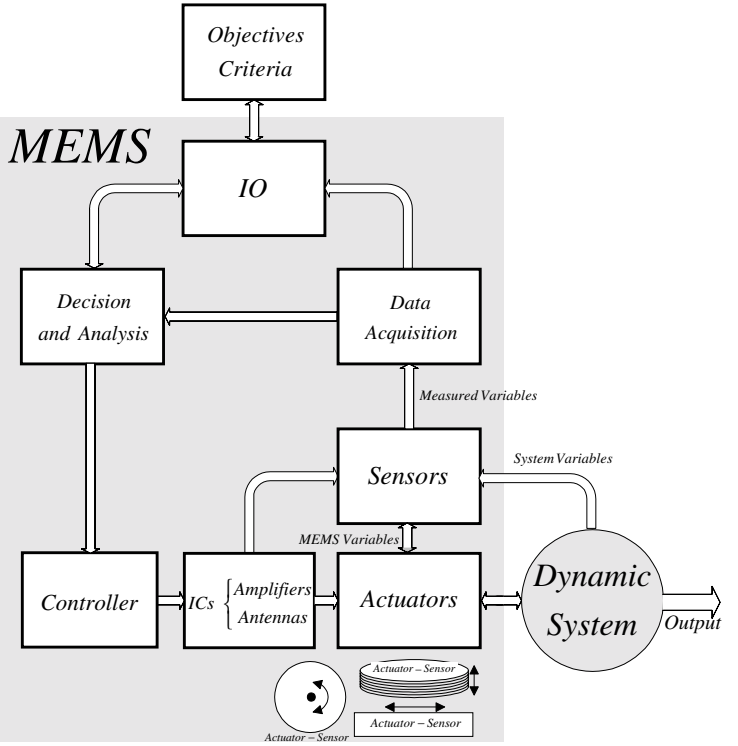


Figure 1.4.2. High-level functional block diagram of large-scale MEMS with rotational and translational actuators and sensors

Actuators are needed to actuate dynamic systems. Actuators respond to command stimulus (control signals) and develop torque and force. There is a great number of biological (e.g., human eye and locomotion system) and man-made actuators. Biological actuators are based upon electromagnetic-mechanical-chemical phenomena and processes. Man-made actuators (electromagnetic, electric, hydraulic, thermo, and acoustic motors) are devices that receive signals or stimulus (stress or pressure, thermo or acoustic, et cetera) and respond with torque or force.

Consider the flight vehicles. The aircraft, spacecraft, missiles, and interceptors are controlled by displacing the control surfaces as well as by changing the control surface and wing geometry. For example, ailerons, elevators, canards, flaps, rudders, stabilizers and tips of advanced aircraft can be controlled by nano-, micro-, and miniscale actuators using the NEMS- and

MEMS-based smart actuator technology. This NEMS- and MEMS-based smart actuator technology is uniquely suitable in the flight actuator applications. Figure 1.4.3 illustrates the aircraft where translational and rotational actuators are used to actuate the control surfaces, as well as to change the wing and control surface geometry.

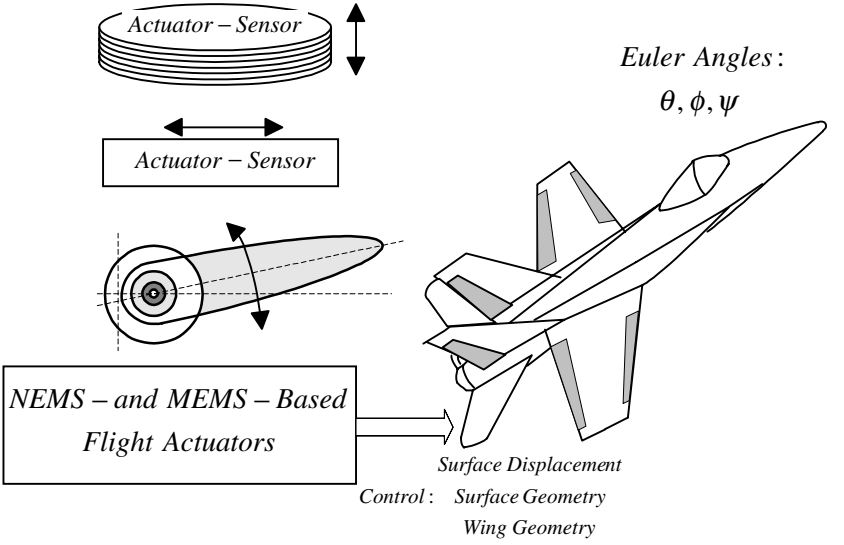


Figure 1.4.3. Aircraft with NEMS- and MEMS-based translational and rotational flight actuators

Sensors are devices that receive and respond to signals or stimulus. For example, the loads (which the aircraft experience during the flight), vibrations, temperature, pressure, velocity, acceleration, noise, and radiation can be measured by micro- and nanoscale sensors, see Figure 1.4.4. It should be emphasized that there are many other sensors to measure the electromagnetic interference and displacement, orientation and position, voltages and currents in power electronic devices, et cetera.

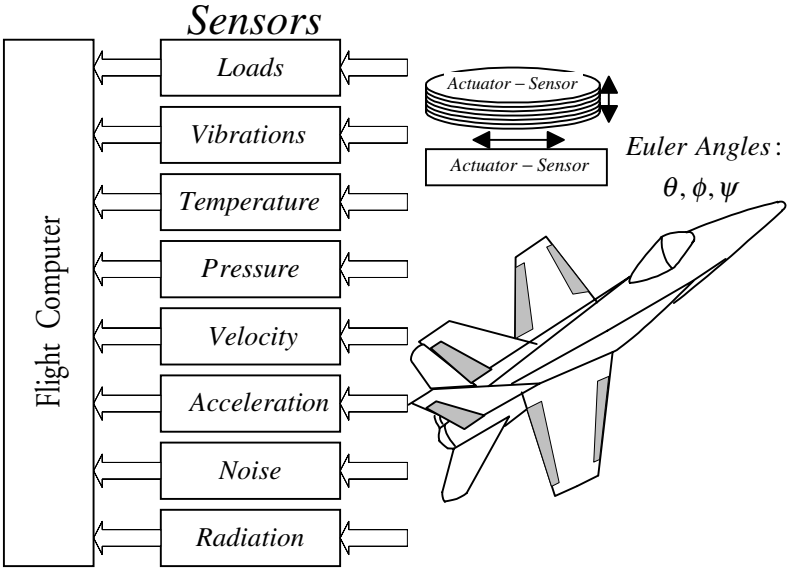


Figure 1.4.4. Application of nano- and microscale sensors in aircraft

Usually, several conversion processes are involved to produce electric, electromagnetic, or mechanical output sensor signals. The conversion of energy is our particular interest. Using the energy-based analysis, the general theoretical fundamentals will be thoroughly studied.

The major developments in NEMS and MEMS have been fabrication technology driven, and the applied research has been performed mainly to manufacture structures and devices, as well as to analyze some performance characteristics. For example, mini- and microscale smart structures as well as ICs have been studied in details, and feasible manufacturing technologies, materials, and processes have been developed. Recently, carbon nanotubes were discovered, and molecular wires and molecular transistors were built. However, to our best knowledge, nanostructures and nanodevices, NEMS and MEMS, have not been comprehensively studied at the nanoscale, and the efforts to develop the fundamental theory have not been reported. In this book, we will apply the quantum theory and charge density concept, advanced electromechanics and Maxwell's equations, as well as other cornerstone methods, to model nanostructures and nanodevices (ICs and antennas, actuators and sensors, et cetera). In particular, the nanoelectromechanical theory will be developed. A large variety of actuators and sensors, antennas and ICs with different operating features are modeled and simulated. To perform high-fidelity integrated 3D+ data intensive modeling with post-processing and animation, the partial and ordinary nonlinear differential equations are solved.

1.5. NANO- AND MICROELECTROMECHANICAL SYSTEMS

In general, monolithic MEMS are integrated microassembled structures (electromechanical microsystems on a single chip) that have both electrical-electronic (ICs) and mechanical components. To manufacture MEMS, advanced modified microelectronics fabrication techniques, technologies, and materials are used. Actuation and sensing cannot be viewed as the peripheral function in many applications. Integrated sensors-actuators (usually motion microstructures) with ICs compose the major class of MEMS. Due to the use of CMOS lithography-based technologies in fabrication actuators and sensors, MEMS leverage microelectronics in important additional areas that revolutionize the application capabilities. In fact, MEMS have considerably leveraged the microelectronics industry beyond ICs. The needs for augmented motion microstructures (actuators and sensors) and ICs have been widely recognized. Simply scaling conventional electromechanical motion devices and augmenting them with ICs have not met the needs, and theory and fabrication processes have been developed beyond component replacement. Dual power operational amplifiers (e.g., Motorola TCA0372, DW Suffix plastic package case 751G, DP2 Suffix plastic package case 648 or DP1 Suffix plastic package case 626) as monolithic ICs can be used to control DC micro electric machines (motion microstructures), as shown in Figure 1.5.1.

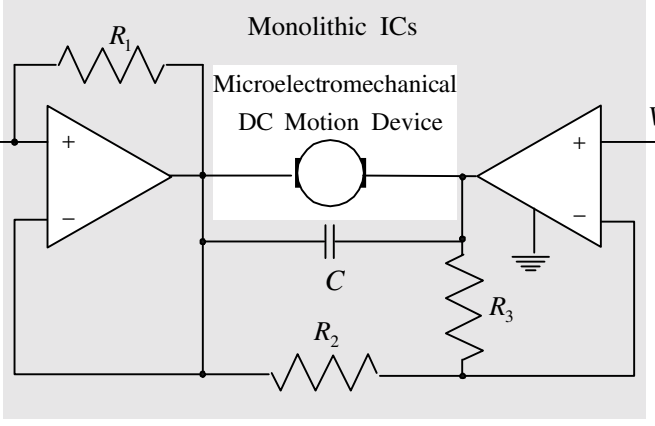


Figure 1.5.1. Application of monolithic IC to control DC micromachines (motion microstructures)

Only recently has it become possible to manufacture MEMS at low cost. However, there is a critical demand for continuous fundamental, applied, and technological improvements, and multidisciplinary activities are required. The general lack of synergy theory to augment actuation, sensing, signal processing, and control is known, and these issues must be addressed through

focussed efforts. The set of long-range goals that challenge the analysis, design, development, fabrication, and deployment of high-performance MEMS are:

- advanced materials and process technology,
- microsensors and microactuators (motion microstructures), sensing and actuation mechanisms, sensors-actuators-ICs integration and MEMS configurations,
- fabrication, packaging, microassembly, and testing,
- MEMS analysis, design, optimization, and modeling,
- MEMS applications and their deployment.

Significant progress in the application of CMOS technology enables the industry to fabricate microscale actuators and sensors with the corresponding ICs, and this guarantees the significant breakthrough. The field of MEMS has been driven by the rapid global progress in ICs, VLSI, solid-state devices, materials, microprocessors, memories, and DSPs that have revolutionized instrumentation, control, and systems design philosophy. In addition, this progress has facilitated explosive growth in data processing and communications in high-performance systems. In microelectronics, many emerging problems deal with nonelectric effects, phenomena and processes (thermal and structural analysis and optimization, stress and ruggedness, packaging, et cetera). It has been emphasized that ICs are the necessary components to perform control, data acquisition, and decision making. For example, control signals (voltage or currents) are computed, converted, modulated, and fed to actuators. It is evident that MEMS have found applications in a wide array of microscale devices (accelerometers, pressure sensors, gyroscopes, et cetera) due to extremely-high level of integration of electromechanical components with low cost and maintenance, accuracy, efficiency, reliability, ruggedness, and survivability. Microelectronics with integrated sensors and actuators are batch-fabricated as integrated assemblies.

Therefore, MEMS can be defined as *batch-fabricated microscale devices (ICs and motion microstructures) that convert physical parameters to electrical signals and vice versa, and in addition, microscale features of mechanical and electrical components, architectures, structures, and parameters are important elements of their operation and design.*

The manufacturability issues in NEMS and MEMS must be addressed. One can design and manufacture individually-fabricated devices and subsystems (ICs and motion microstructures). However, these individually-fabricated devices and subsystems are unlikely can be used due to very high cost.

Integrated MEMS combine mechanical structures (microfabricated smart multifunctional materials are used to manufacture microscale actuators and sensors, pumps and valves, optical devices) and microelectronics (ICs). The number of transistors on a chip is frequently used by the microelectronic industry, and enormous progress in achieving nanoscale transistor dimensions

(less than 100 nm) was achieved. However, large-scale MEMS operational capabilities are measured by the intelligence, system-on-a-chip integration, integrity, cost, performance, efficiency, size, reliability, and other criteria. There are a number of challenges in MEMS fabrication because conventional CMOS technology must be modified and integration strategies (to integrate mechanical structures and ICs) are needed to be developed. What (ICs or mechanical micromachined structure) should be fabricated first? Fabrication of ICs first faces challenges because to reduce stress in the thin films of polysilicon (multifunctional material to build motion microstructures), a high-temperature anneal at 1000^0C is needed for several hours. The aluminum ICs interconnect will be destroyed (melted), and tungsten can be used for interconnected metallization. This process leads to difficulties for commercially manufactured MEMS due to high cost and low reproducibility. Analog Devices fabricates ICs first up to metallization step, and then, mechanical structures (polysilicon) are built using high-temperature anneal (micromachines are fabricated before metallization), and finally, ICs are interconnected. This allows the manufacturer to use low-cost conventional aluminum interconnects. The third option is to fabricate mechanical structures, and then ICs. However, to overcome step coverage, stringer, and topography problems, motion mechanical microstructures can be fabricated in the bottoms of the etched shallow trenches (packaged directly) of the wafer. These trenches are filled with a sacrificial silicon dioxide, and the silicon wafer is planarized through chemical-mechanical polishing.

The motion mechanical microstructures can be protected (sensor applications, e.g., accelerometers and gyroscopes) and unprotected (actuator and interactive environment sensor applications). Therefore, MEMS (mechanical structure – ICs) can be encased in a clean, hermetically sealed package or some elements can be unprotected to interact with environment. This creates challenges in packaging. It is extremely important to develop novel electromechanical motion microstructures and microdevices (sticky multilayers, thin films, magnetoelectronic, electrostatic, and quantum-effect-based devices) and sense their properties. Microfabrication of very large scale integrated circuits (VLSI), MEMS, and optoelectronics must be addressed. Fabrication processes include lithography, film growth, diffusion, ion implantation, thin film deposition, etching, metallization, et cetera. Furthermore, ICs and motion microstructures (microelectromechanical motion devices) must be connected. Complete microfabrication processes with integrated process steps must be developed.

Microelectromechanical systems integrate microscale subsystems (at least ICs and motion structure). It was emphasized that microsensors sense the physical variables, and microactuators control (actuate) real-world systems. These microactuators are regulated by ICs. It must be emphasized that ICs also performed computations, signal conditioning, decision making, and other

functions. For example, in microaccelerometers, the motion microstructure displaces. Using this displacement, the acceleration can be calculated. In microaccelerometers, computations, signal conditioning, data acquisition, and decision making are performed by ICs. Microactuators inflate air-bags if car crashes (high g acceleration measured).

Microelectromechanical systems contain microscale subsystems designed and manufactured using different technologies. Single silicon substrate can be used to fabricate microscale actuators, sensors, and ICs (monolithic MEMS) using CMOS microfabrication technology. Alternatively, subsystem can be assembled, connected and packaged, and different microfabrication techniques for MEMS components and subsystems exist. Usually, monolithic MEMS are compact, efficient, reliable, and guarantee superior performance.

Typically, MEMS integrate the following subsystems: microscale actuators (actuate real-world systems), microscale sensors (detect and measure changes of the physical variables), and microelectronics/ICs (signal processing, data acquisition, decision making, et cetera).

Microactuators are needed to develop force or torque (mechanical variable). Typical examples are microscale drives, moving mirrors, pumps, servos, valves, et cetera. A great variety of methods for achieving actuation are well-known, e.g., electromagnetic (electrostatic, magnetic, piezoelectric), hydraulic, and thermal effects. This book covers electromagnetic microactuators, and the so-called comb drives (surface micromachined motion microstructures) have been widely used. These drives have movable and stationary plates (fingers). When the voltage is applied, an attractive force is developed between two plates, and the motion results. A wide variety of microscale actuators have been fabricated and tested. The common problem is the difficulties associated with coil fabrication. The choice of magnetic materials (permanent magnets) is limited to those that can be micromachined. Magnetic actuators typically fabricated through the photolithography technology using nickel (ferromagnetic material). Piezoelectric microactuators have found wide applications due to simplicity and ruggedness (force is generated if one applies the voltage across a film of piezoelectric material). The piezoelectric-based concept can be applied to thin silicon membranes, and if the voltage is applied, the membrane deforms. Thus, silicon membranes can be used as pumps.

Microsensors are devices that convert one physical variable (quantity) to another. For example, electromagnetic phenomenon can be converted to mechanical or optic effects. There are a number of different types of microscale sensors used in MEMS. For example, microscale thermosensors are designed and built using the thermoelectric effect (the resistivity varies with temperature). Extremely low cost thermoresistors (thermistors) are fabricated on the silicon wafer, and ICs are built on the same substrate. The thermistor resistivity is a highly nonlinear function of the temperature, and the compensating circuitry is used to take into account the nonlinear effect. Microelectromagnetic sensors measure electromagnetic fields, e.g., the Hall

effect sensors. Optical sensors can be fabricated on crystals that exhibit a magneto-optic effect, e.g., optical fibers. In contrast, the quantum effect sensors can sense extremely weak electromagnetic fields. Silicon-fabricated piezoresistors (silicon doped with impurities to make it n- or p-type) belong to the class of mechanical sensors. When the force is applied to the piezoelectric, the charge induced (measured voltage) is proportional to the applied force. Zinc oxide and lead zirconate titanate (PZT, $PbZrTiO_3$), which can be deposited on microstructures, are used as piezoelectric crystals. In this book, the microscale accelerometers and gyroscopes, as well as microelectric machines will be studied. Accelerometers and gyroscopes are based upon capacitive sensors. In two parallel conducting plates, separated by an insulating material, the capacitance between the plates is a function of distance between plates (capacitance is inversely proportional to the distance). Thus, measuring the capacitance, the distance can be easily calculated. In accelerometers and gyroscopes, the proof mass and rotor are suspended. It will be shown that using the second Newton's law, the acceleration is proportional to the displacement. Hence, the acceleration can be calculated. Thin membranes are the basic components of pressure sensors. The deformation of the membrane is usually sensed by piezoresistors or capacitive microsensors.

We have illustrated the critical need for physical- and system-level concepts in NEMS and MEMS analysis and design. Advances in physical-level research have tremendously expanded the horizon of NEMS and MEMS technologies. For example, magnetic-based (magnetoelectronic) memories have been thoroughly studied (magnetoelectronic devices are grouped in three categories based upon the physics of their operation: all-metal spin transistors and valves, hybrid ferromagnetic semiconductor structures, and magnetic tunnel junctions). Writing and reading the cell data are based on different physical mechanisms, and high or low cost, densities, power, reliability and speed (write/read cycle) memories result. As the physical-level analysis and design are performed, the system-level analysis and design must be accomplished because the design of integrated large-scale NEMS and MEMS is the final goal.

1.6. INTRODUCTION TO MEMS FABRICATION, ASSEMBLING, AND PACKAGING

Two basic components of MEMS and microengineering are microelectronics (to fabricate ICs) and micromachining (to fabricate motion microstructures). Using CMOS or VLSI technology, microelectronics (ICs) fabrication can be performed. Micromachining technology is needed to fabricate motion microstructures to be used as the MEMS mechanical subsystems. It was emphasized that one of the main goals of microengineering is to integrate microelectronics with micromachined

mechanical structures in order to produce completely integrated monolithic high-performance MEMS. To guarantee low cost, reliability, and manufacturability, the following must by guaranteed: the fabrication process has a high yield and batch processing techniques are used for as much of the process as possible (large numbers of microscale structures/devices per silicon wafer and large number of wafers are processed at the same time at each fabrication step). Assembling and packaging must be automated, and the most promising avenues are auto- or self-alignment and self assembly. Some MEMS subsystems (actuator and interactive environment sensors) must be protected from mechanical damage, and in addition, protected from contamination. Wear tolerance, electromagnetic and thermo isolation, among other problems have always challenged MEMS. Different manufacturing technologies must be applied to attain the desired performance level and cost. Microsubsystems can be coated directly by thin films of silicon dioxide or silicon nitride which are deposited using plasma enhanced chemical vapor deposition. It is possible to deposit (at 700^0C to 900^0C) films of diamond which have superior wear capabilities, excellent electric insulation and thermal characteristics. It must be emphasized that diamond like carbon films can be also deposited.

Microelectromechanical systems are connected (interfaced) with real-world systems (control surfaces of aircraft, flight computer, communication ports, et cetera). Furthermore, MEMS are packaged to protect systems from harsh environments, prevent mechanical damage, minimize stresses and vibrations, contamination, electromagnetic interference, et cetera. Therefore, MEMS are usually sealed. It is impossible to specify a generic MEMS package. Through input-output connections (power and communication bus) one delivers the power required, feeds control (command) and test (probe) signals, receives the output signals and data. Packages must be designed to minimize electromagnetic interference and noise. Heat, generated by MEMS, must be dissipated, and the thermal expansion problem must be solved. Conventional MEMS packages are usually ceramic and plastic. In ceramic packages, the die is bonded to a ceramic base, which includes a metal frame and pins for making electric outside connections. Plastic packages are connected in the similar way. However, the package can be molded around the microdevice.

Silicon and silicon carbide micromachining are the most developed micromachining technologies. Silicon is the primary substrate material which is used by the microelectronics industry. A single crystal ingot (solid cylinder 300 mm diameter and 1000 mm length) of very high purity silicon is grown, then sawed with the desired thickness and polished using chemical and mechanical polishing techniques. Electromagnetic and mechanical wafer properties depend upon the orientation of the crystal growth, concentration and type of doped impurities. Depending on the silicon substrate, CMOS processes are used to manufacture ICs, and the process is classified as *n-well*, *p-well*, or *twin-well*. The major steps are diffusion, oxidation, polysilicon gate formations, photolithography, masking, etching, metallization, wire bonding, et cetera. To fabricate motion microstructures (microelectromechanical motion devices),

CMOS technology must be modified. High-resolution photolithography is a technology that is applied to produce moulds for the fabrication of micromachined mechanical components and to define their three-dimensional shape (geometry). That is, the micromachine geometry is defined photographically. First, a mask is produced on a glass plate. The silicon wafer is then coated with a polymer which is sensitive to ultraviolet light (photoresistive layer is called photoresist). Ultraviolet light is shone through the mask onto the photoresist to build the mask to the photoresist layer. The positive photoresist becomes softened, and the exposed layer can be removed. In general, there are two types of photoresist, e.g., positive and negative. Where the ultraviolet light strikes the positive photoresist, it weakens the polymer. Hence, when the image is developed, the photoresist is washed where the light struck it. A high-resolution positive image results. In contrast, if the ultraviolet light strikes negative photoresist, it strengthens the polymer. Therefore, a negative image of the mask results. Chemical process is used to remove the oxide where it is exposed through the openings in the photoresist. When the photoresist is removed, the patterned oxide appears. Alternatively, electron beam lithography can be used. Photolithography requires design of masks. The design of photolithography masks for micromachining is straightforward, and computer-aided-design (CAD) software is available and widely applied.

There are a number of basic surface silicon micromachining technologies that can be used in order to pattern thin films that have been deposited on a silicon wafer, and to shape the silicon wafer itself forming a set of basic microstructures. Three basic steps associated with silicon micromachining are:

- deposition of thin films of materials;
- removal of material (patterning) by wet or dry techniques;
- doping.

Different microelectromechanical motion devices (motion microstructures) can be designed, and silicon wafers with different crystal orientations are used. Reactive ion etching (dry etching) is usually applied. Ions are accelerated towards the material to be etched, and the etching reaction is enhanced in the direction of ion traveling. Deep trenches and pits of desired shapes can be etched in a variety of materials including silicon, oxide, and nitride. A combination of dry and wet etching can be embedded in the process.

Metal films are patterned using the lift off stenciling technique. A thin film of the assisting material (oxide) is deposited, and a layer of photoresist is put over and patterned. The oxide is then etched to undercut the photoresist. The metal film is then deposited on the silicon wafer through evaporation process. The metal pattern is stenciled through the gaps in the photoresist, which is then removed, lifting off the unwanted metal. The assisting layer is then stripped off, leaving the metal film pattern.

The anisotropic wet etching and concentration dependent etching are

called bulk silicon micromachining because the microstructures are formed by etching away the bulk of the silicon wafer. Surface micromachining forms the structure in layers of thin films on the surface of the silicon wafer or other substrate. Hence, the surface micromachining process uses thin films of two different materials, e.g., structural (usually polysilicon) and sacrificial (oxide) materials. Sacrificial layers of oxide are deposited on the wafer surface, and dry etched. Then, the sacrificial material is wet etched away to release the structure. A variety of different complex motion microstructures with different geometry have been fabricated using the surface micromachining technology.

Micromachined silicon wafers must be bonded together. Anodic (electrostatic) bonding technique is used to bond silicon wafer and glass substrate. In particular, the silicon wafer and glass substrate are attached, heated, and electric field is applied across the join. These result in extremely strong bonds between the silicon wafer and glass substrate. In contrast, the direct silicon bonding is based upon applying pressure to bond silicon wafer and glass substrate. It must be emphasized that to guarantee strong bonds, the silicon wafer and glass substrate surfaces must be flat and clean.

The MEMCAD™ software (current version is 4.6), developed by Microcosm, is widely used to design, model, simulate, characterize, and package MEMS. Using the built-in Microcosm Catapult™ layout editor, augmented with materials database and components library, three-dimensional solid models of motion microstructures can be developed. Furthermore, customizable packaging is fully supported.

CHAPTER 2

MATHEMATICAL MODELS AND DESIGN OF NANO- AND MICROELECTROMECHANICAL SYSTEMS

2.1. NANO- AND MICROELECTROMECHANICAL SYSTEMS ARCHITECTURE

A large variety of nano- and microscale structures and devices, as well as NEMS and MEMS (systems integrate structures, devices, and subsystems), have been widely used, and a worldwide market for NEMS and MEMS and their applications will be drastically increased in the near future. The differences in NEMS and MEMS are emphasized, and NEMS are smaller than MEMS. For example, carbon nanotubes (nanostructure) can be used as the molecular wires and sensors in MEMS. Different specifications are imposed on NEMS and MEMS depending upon their applications. For example, using carbon nanotubes as the molecular wires, the current density is defined by the media properties (e.g., resistivity and thermal conductivity). It is evident that the maximum current is defined by the diameter and the number of layers of the carbon nanotube. Different molecular-scale nanotechnologies are applied to manufacture NEMS (controlling and changing the properties of nanostructures), while analog, discrete, and hybrid MEMS have been mainly manufactured using surface micro-machining, silicon-based technology (lithographic processes are used to fabricate CMOS ICs). To deploy and commercialize NEMS and MEMS, a spectrum of problems must be solved, and a portfolio of software design tools needs to be developed using a multidisciplinary concept. In recent years much attention has been given to MEMS fabrication and manufacturing, structural design and optimization of actuators and sensors, modeling, analysis, and optimization. It is evident that NEMS and MEMS can be studied with different level of detail and comprehensiveness, and different application-specific architectures should be synthesized and optimized. The majority of research papers study either nano- and microscale actuators-sensors or ICs that can be the subsystems of NEMS and MEMS. A great number of publications have been devoted to the carbon nanotubes (nanostructures used in NEMS and MEMS). The results for different NEMS and MEMS components are extremely important and manageable. However, the comprehensive systems-level research must be performed because the specifications are imposed on the systems, not on the individual elements, structures, and subsystems of NEMS and MEMS. Thus, NEMS and MEMS must be developed and studied to attain the comprehensiveness of the analysis and design.

For example, the actuators are controlled changing the voltage or current (by ICs) or the electromagnetic field (by nano- or microscale antennas). The

ICs and antennas (which should be studied as the subsystems) can be controlled using nano or micro decision-making systems, which can include central processor and memories (as core), IO devices, etc. Nano- and microscale sensors are also integrated as elements of NEMS and MEMS, and through molecular wires (for example, carbon nanotubes) one feeds the information to the IO devices of the nano-processor. That is, NEMS and MEMS integrate a large number of structures and subsystems which must be studied. As a result, the designer usually cannot consider NEMS and MEMS as six-degrees-of-freedom actuators using conventional mechanics (the linear or angular displacement is a function of the applied force or torque), completely ignoring the problem of how these forces or torques are generated and regulated. In this book, we will illustrate how to integrate and study the basic components of NEMS and MEMS.

The design and development, modeling and simulation, analysis and prototyping of NEMS and MEMS must be attacked using advanced theories. The systems analysis of NEMS and MEMS as systems integrates analysis and design of structures, devices and subsystems used, structural optimization and modeling, synthesis and optimization of architectures, simulation and virtual prototyping, etc. Even though a wide range of nanoscale structures and devices (e.g., molecular diodes and transistors, machines and transducers) can be fabricated with atomic precision, comprehensive systems analysis of NEMS and MEMS must be performed before the designer embarks in costly fabrication because through optimization of architecture, structural optimization of subsystems (actuators and sensors, ICs and antennas), modeling and simulation, analysis and visualization, the rapid evaluation and prototyping can be performed facilitating cost-effective solution reducing the design cycle and cost, guaranteeing design of high-performance NEMS and MEMS which satisfy the requirements and specifications.

The large-scale integrated MEMS (a single chip that can be mass-produced using the CMOS, lithography, and other technologies at low cost) integrates:

- N nodes of actuators/sensors, smart structures, and antennas;
- processor and memories,
- interconnected networks (communication busses),
- input-output (IO) devices,
- etc.

Different architectures can be implemented, for example, linear, star, ring, and hypercube are illustrated in Figure 2.1.1.

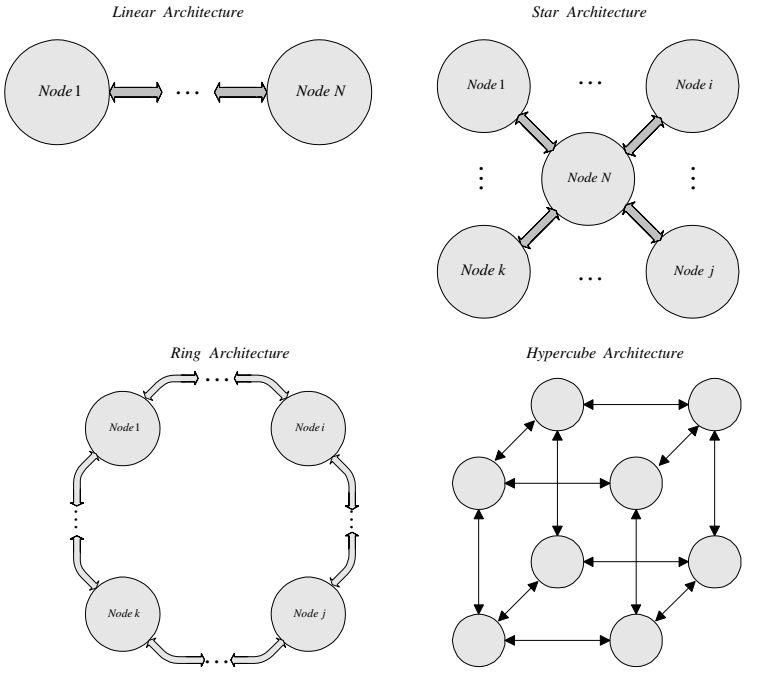


Figure 2.1.1. Linear, star, ring, and hypercube architectures

More complex architectures can be designed, and the hypercube-connected-cycle node configuration is illustrated in Figure 2.1.2.

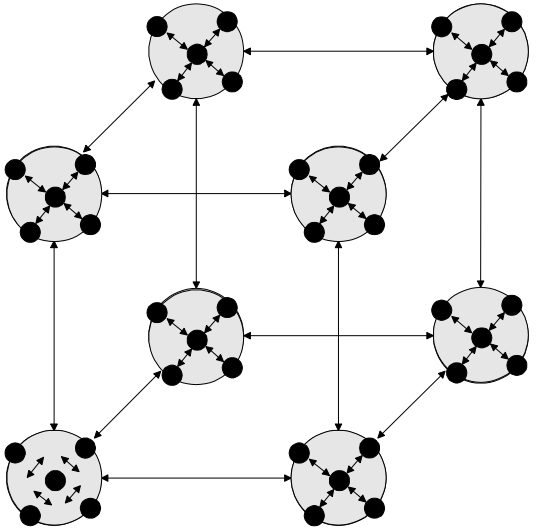


Figure 2.1.2. Hypercube-connected-cycle node architecture

The nodes can be synthesized, and the elementary node can be simply pure smart structure, actuator, or sensor. This elementary node can be controlled by the external electromagnetic field (that is, ICs or antenna are not a part of the elementary structure). In contrast, the large-scale node can integrate processor (with decision making, control, signal processing, and data acquisition capabilities), memories, IO devices, communication bus, ICs and antennas, actuators and sensors, smart structures, etc. That is, in addition to actuators/sensors and smart structures, ICs and antennas (to regulate actuators/sensors and smart structures), processor (to control ICs and antennas), memories and interconnected networks, IO devices, as well as other subsystems can be integrated. Figure 2.1.3 illustrates large-scale and elementary nodes.

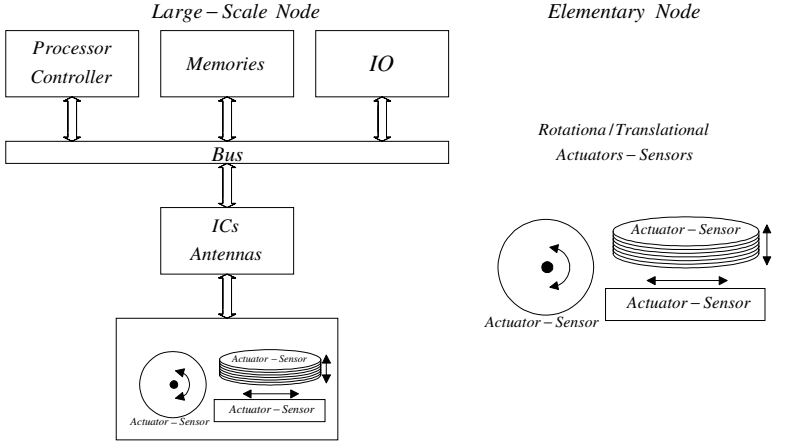


Figure 2.1.3. Large-scale and elementary nodes

As NEMS and MEMS are used to control physical dynamic systems (immune system or drug delivery, propeller or wing, relay or lock), to illustrate the basic components, a high-level functional block diagram is shown in Figure 2.1.4.

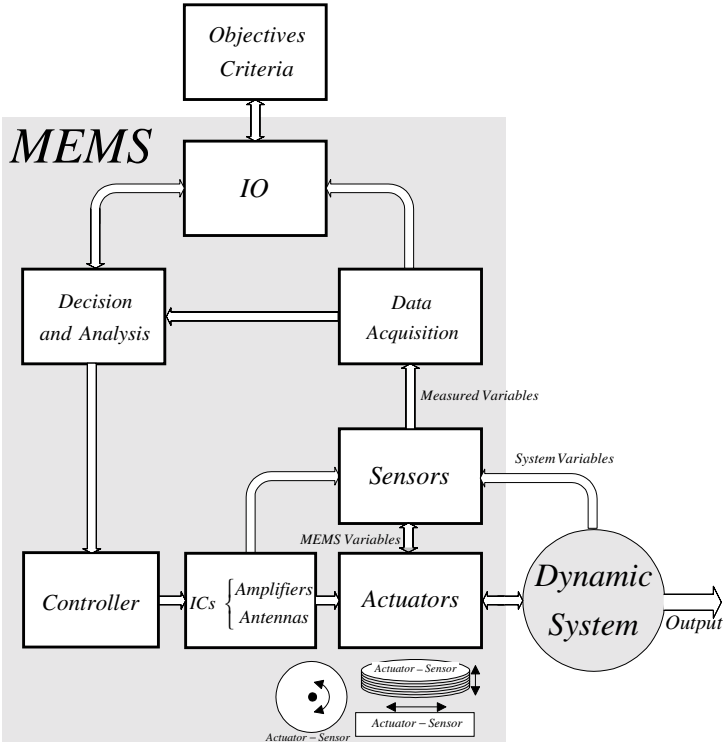


Figure 2.1.4. High-level functional block diagram of large-scale NEMS and MEMS

For example, the desired flight path of aircraft (maneuvering and landing) is maintained by displacing the control surfaces (ailerons and elevators, canards and flaps, rudders and stabilizers) and/or changing the control surface and wing geometry. Figure 2.1.5 documents the application of the NEMS- and MEMS-based technology to actuate the control surfaces. It should be emphasized that the NEMS and MEMS receive the digital signal-level signals from the flight computer, and these digital signals are converted into the desired voltages or currents fed to the microactuators or electromagnetic flux intensity to displace the actuators. It is also important that NEMS- and MEMS-based transducers can be used as sensors, and, as an example, the loads on the aircraft structures during the flight can be measured.

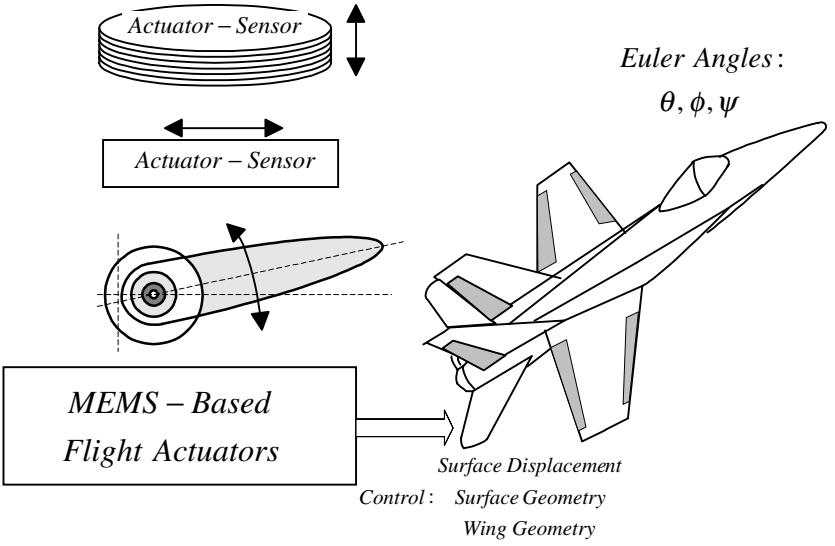


Figure 2.1.5. Aircraft with MEMS-based flight actuators

Microelectromechanical and Nanoelectromechanical Systems

Microelectromechanical systems are integrated microassembled structures (electromechanical microsystems on a single chip) that have both electrical-electronic (ICs) and mechanical components. To manufacture MEMS, modified advanced microelectronics fabrication techniques and materials are used. It was emphasized that sensing and actuation cannot be viewed as the peripheral function in many applications. Integrated actuators/sensors with ICs compose the major class of MEMS. Due to the use of CMOS lithography-based technologies in fabrication actuators and sensors, MEMS leverage microelectronics (signal processing, computing, and control) in important additional areas that revolutionize the application capabilities. In fact, MEMS have been considerably leveraged the microelectronics industry beyond ICs. The needs to augmented actuators, sensors, and ICs have been widely recognized. For example, mechatronics concept, used for years in conventional electromechanical systems, integrates all components and subsystems (electromechanical motion devices, power converters, microcontrollers, et cetera). Simply scaling conventional electromechanical motion devices and augmenting them with ICs have not

met the needs, and theory and fabrication processes have been developed beyond component replacement. Only recently it becomes possible to manufacture MEMS at very low cost. However, there is a critical demand for continuous fundamental, applied, and technological improvements, and multidisciplinary activities are required. The general lack of synergy theory to augment actuation, sensing, signal processing, and control is known, and these issues must be addressed through focussed efforts. The set of long-range goals has been emphasized in Chapter 1. The challenges facing the development of MEMS are

- advanced materials and process technology,
- microsensors and microactuators, sensing and actuation mechanisms, sensors-actuators-ICs integration and MEMS configurations,
- packaging, microassembly, and testing,
- MEMS modeling, analysis, optimization, and design,
- MEMS applications and their deployment.

Significant progress in the application of CMOS technology enable the industry to fabricate microscale actuators and sensors with the corresponding ICs, and this guarantees the significant breakthrough. The field of MEMS has been driven by the rapid global progress in ICs, VLSI, solid-state devices, microprocessors, memories, and DSPs that have revolutionized instrumentation and control. In addition, this progress has facilitated explosive growth in data processing and communications in high-performance systems. In microelectronics, many emerging problems deal with nonelectric phenomena and processes (thermal and structural analysis and optimization, packaging, et cetera). It has been emphasized that ICs is the necessary component to perform control, data acquisition, and decision making. For example, control signals (voltage or currents) are computer, converted, modulated, and fed to actuators. It is evident that MEMS have found application in a wide array of microscale devices (accelerometers, pressure sensors, gyroscopes, et cetera) due to extremely-high level of integration of electromechanical components with low cost and maintenance, accuracy, reliability, and ruggedness. Microelectronics with integrated sensors and actuators are batch-fabricated as integrated assemblies.

Therefore, MEMS can be defined as

batch-fabricated microscale devices (ICs and motion microstructures) that convert physical parameters to electrical signals and vise versa, and in addition, microscale features of mechanical and electrical components, architectures, structures, and parameters are important elements of their operation and design.

The manufacturability issues in NEMS and MEMS must be addressed. It was shown that one can design and manufacture individually-fabricated devices and subsystems. However, these devices and subsystems are unlikely will be used due to very high cost.

Piezoactuators and permanent-magnet technology has been used widely, and rotating and linear electric transducers (actuators and sensors) are designed. For example, piezoactive materials are used in ultrasonic motors. Frequently, conventional concepts of the electric machinery theory (rotational and linear direct-current, induction, and synchronous machine) are used to design and analyze MEMS-based machines. The use of piezoactuators is possible as a consequence of the discovery of advanced materials in sheet and thin-film forms, especially PZT (lead zirconate titanate) and polyvinylidene fluoride. The deposition of thin films allows piezo-based electric machines to become a promising candidate for microactuation in lithography-based fabrication. In particular, microelectric machines can be fabricated using a deep x-ray lithography and electrodeposition process. Two-pole synchronous and induction micro-motors have been fabricated and tested.

To fabricate nanoscale structures, devices, and NEMS, molecular manufacturing methods and technologies must be developed. Self- and positional-assembly concepts are the preferable technologies compared with individually-fabricated in the synthesis and manufacturing of molecular structures. To perform self- and positional-assembly, complementary pairs (CP) and molecular building blocks (MBB) should be designed. These CP or MBB, which can be built from a couple to thousands atoms, can be studied and designed using the DNA analogy. The nucleic acids consist of two major classes of molecules (DNA and RNA). Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the largest and most complex organic molecules which are composed of carbon, oxygen, hydrogen, nitrogen, and phosphorus. The structural units of DNA and RNA are nucleotides, and each nucleotide consists of three components (nitrogen-base, pentose and phosphate) joined by dehydration synthesis. The double-helix molecular model of DNA was discovered by Watson and Crick in 1953. The DNA (long double-stranded polymer with double chain of nucleotides held together by hydrogen bonds between the bases), as the genetic material (genes), performs two fundamental roles. It replicates (identically reproduces) itself before a cell divides, and provides pattern for protein synthesis directing the growth and development of all living organisms according to the information DNA supports. The DNA architecture provides the mechanism for the replication of genes. Specific pairing of nitrogenous bases obey base-pairing rules and determine the combinations of nitrogenous bases that form the rungs of the double helix. In contrast, RNA carries (performs) the protein synthesis using the DNA information. Four DNA bases are: A (adenine), G (guanine), C (cytosine), and T (thymine). The ladder-like DNA molecule is formed due to hydrogen bonds between the bases which paired in the interior of the double helix (the base pairs are 0.34 nm apart and there are ten pairs per turn of the helix). Two backbones (sugar and phosphate molecules) form the uprights of the DNA molecule, while the joined bases form the rungs.

Figure 2.1.6 illustrates that the hydrogen bonding of the bases are: A bonds to T, G bonds to C. The complementary base sequence results.

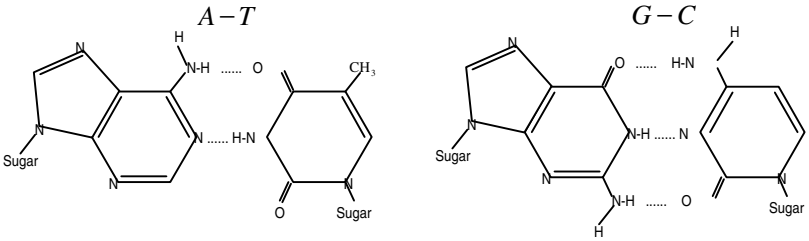


Figure 2.1.6. DNA pairing due to hydrogen bonds

In RNA molecules (single strands of nucleotides), the complementary bases are A bonds to U (uracil), and G bonds to C. The complementary base bonding of DNA and RNA molecules gives one the idea of possible sticky-ended assembling (through complementary pairing) of NEMS structures and devices with the desired level of specificity, architecture, topology, and organization. In structural assembling and design, the key element is the ability of CP or MBB (atoms or molecules) to associate with each other (recognize and identify other atoms or molecules by means of specific base pairing relationships). It was emphasized that in DNA, A (adenine) bonds to T (thymine) and G (guanine) bonds to C (cytosine). Using this idea, one can design the CP such as A_1-A_2, B_1-B_2, C_1-C_2, etc. That is, A_1 pairs with A_2, while B_1 pairs with B_2. This complementary pairing can be studied using electromagnetics (Coulomb law) and chemistry (chemical bonding, for example, hydrogen bonds in DNA between nitrogenous bases A and T, G and C). Figure 2.1.7 shows how two nanoscale elements with sticky ends form the complementary pair. In particular, "+" is the sticky end and "-" is its complement. That is, the complementary pair A_1-A_2 results.

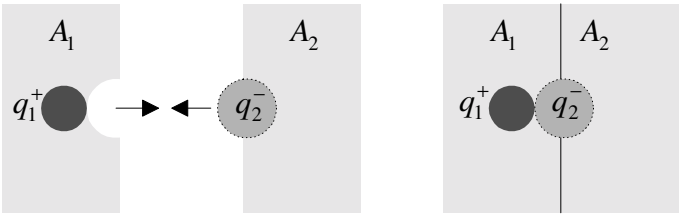


Figure 2.1.7. Sticky ended electrostatically complementary pair A_1-A_2

An example of assembling a ring is illustrated in Figure 2.1.8. Using the sticky ended segmented (asymmetric) electrostatically CP, self-assembling of

nanostructure is performed in the XY plane. It is evident that three-dimensional structures can be formed through the self-assembling.

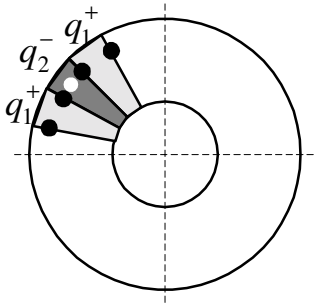


Figure 2.1.8. Ring self-assembling

It is evident that there are several advantages to use sticky ended electrostatic CP. In the first place, the ability to recognize (identify) the complementary pair is clear and reliably predicted. The second advantage is the possibility to form stiff, strong, and robust structures.

Self-assembled complex nanostructures can be fabricated using subsegment concept to form the branched junctions. This concept is well-defined electrostatically and geometrically through Coulomb law and branching connectivity. Using the subsegment concept, ideal objects (e.g., cubes, octahedron, spheres, cones, et cetera) can be manufactured. Furthermore, the geometry of nanostructures can be easily controlled by the number of CP and pairing MBB. It must be emphasized that it is possible to generate a quadrilateral self-assembled nanostructure by using four and more different CP. That is, in addition to electrostatic CP, chemical CP can be used. Single- and double-stranded structures can be generated and linked in the desired topological and architectural manners. The self-assembling must be controlled during the manufacturing cycle, and CP and MBB, which can be paired and topologically/architecturally bonded, must be added in the desired sequence. For example, polyhedral and octahedral synthesis can be performed when building elements (CP or MBB) are topologically or geometrically specified. The connectivity of nanostructures determines the minimum number of linkages that flank the branched junctions. The synthesis of complex three-dimensional nanostructures is the design of topology, and the structures are characterized by their branching and linking.

Linkage Groups in Molecular Building Blocks

The hydrogen bonds, which are weak, hold DNA and RNA strands. Strong bonds are desirable to form stiff, strong, and robust nano- and microstructures. Using polymer chemistry, functional groups which couple

monomers can be designed. However, polymers made from monomers with only two linkage groups do not exhibit the desired stiffness and strength. Tetrahedral MBB structures with four linkage groups result in stiff and robust structures. Polymers are made from monomers, and each monomer reacts with two other monomers to form linear chains. Synthetic and organic polymers (large molecules) are nylon and dacron (synthetic), and proteins and RNA, respectively.

There are two major ways to assemble parts. In particular, self assembly and positional assembly. Self-assembling is widely used at the molecular scale, and the DNA and RNA examples were already emphasized. Positional assembling is widely used in manufacturing and microelectronic manufacturing. The current inability to implement positional assembly at the molecular scale with the same flexibility and integrity that it applied in microelectronic fabrication limits the range of nanostructures which can be manufactured. Therefore, the efforts are focused on developments of MBB, as applied to manufacture nanostructures, which guarantee:

- mass-production at low cost and high yield;
- simplicity and predictability of synthesis and manufacturing;
- high-performance, repeatability, and similarity of characteristics;
- stiffness, strength, and robustness;
- tolerance to contaminants.

It is possible to select and synthesize MBB that satisfy the requirements and specifications (non-flammability, non-toxicity, pressure, temperatures, stiffness, strength, robustness, resistivity, permiability, permittivity, et cetera). Molecular building blocks are characterized by the number of linkage groups and bonds. The linkage groups and bonds that can be used to connect MBB are:

- dipolar bonds (weak),
- hydrogen bonds (weak),
- transition metal complexes bonds (weak),
- amide and ester linkages (weak and strong).

It must be emphasized that large molecular building blocks (LMMB) can be made from MBB. There is a need to synthesize robust three-dimensional structures. Molecular building blocks can form planar structures with are strong, stiff, and robust in-plane, but weak and compliant in the third dimension. This problem can be resolved by forming tubular structures. It was emphasized that it is difficult to form three-dimensional structures using MBB with two linkage groups. Molecular building blocks with three linkage groups form planar structures, which are strong, stiff, and robust in plane but bend easily. This plane can be rolled into tubular structures to guarantee stiffness. Molecular building blocks with four, five, six, and twelve linkage groups form strong, stiff, and robust three-dimensional structures needed to synthesize robust nano- and microstructures.

Molecular building blocks with L linkage groups are paired forming L-pair structures, and planar and non-planar (three-dimensional) nano- and

microstructures result. These MBB can have in-plane linkage groups and out-of-plane linkage groups which are normal to the plane. For example, hexagonal sheets are formed using three in-plane linkage groups (MBB is a single carbon atom in a sheet of graphite) with adjacent sheets formed using two out-of-plane linkage groups. It is evident that this structure has hexagonal symmetry.

Molecular building blocks with six linkage groups can be connected together in the cubic structure. These six linkage groups corresponding to six sides of the cube or rhomb. Thus, MBB with six linkage groups form solid three-dimensional structures as cubes or rhomboids. It should be emphasized that buckyballs (C_{60}), which can be used as MMB, are formed with six functional groups. Molecular building blocks with six in-plane linkage groups form strong planar structures. Robust, strong, and stiff cubic or hexagonal closed-packed crystal structures are formed using twelve linkage groups. Molecular building blocks synthesized and applied should guarantee the desirable performance characteristics (stiffness, strength, robustness, resistivity, permiability, permittivity, et cetera) as well as manufacturability. It is evident that stiffness, strength, and robustness are predetermined by bonds (weak and strong), while resistivity, permiability and permittivity are the functions of MBB compounds and media.

2.2. ELECTROMAGNETICS AND ITS APPLICATION FOR NANO- AND MICROSCALE ELECTROMECHANICAL MOTION DEVICES

To study NEMS and MEMS actuators and sensors, smart structures, ICs and antennas, one applies the electromagnetic field theory. Electric force holds atoms and molecules together. Electromagnetics plays a central role in molecular biology. For example, two DNA (deoxyribonucleic acid) chains wrap about one another in the shape of a *double helix*. These two strands are held together by electrostatic forces. Electric force is responsible for energy-transforming processes in all living organisms (*metabolism*). Electromagnetism is used to study protein synthesis and structure, nervous system, etc.

Electrostatic interaction was investigated by Charles Coulomb.

For charges q_1 and q_2, separated by a distance x in free space, the magnitude of the electric force is

$$F = \frac{|q_1 q_2|}{4\pi e_0 x^2},$$

where e_0 is the permittivity of free space, e_0 = 8.85 10 ^{12} F/m or C^2/N-m^2,

$\frac{1}{4\pi e_0} = 9 \quad 10^9$ N-m^2/C.

The unit for the force is the newton N, while the charges are given in coulombs, C.

The force is the vector, and we have

$$\vec{F} = \frac{q_1 q_2}{4\pi e_0 x^2} \vec{a}_x,$$

where $\vec{a}_x$ is the unit vector which is directed along the line joining these two charges.

The capacity, elegance and uniformity of electromagnetics arise from a sequence of fundamental laws linked one to other and needed to study the field quantities.

Using the Gauss law and denoting the vector of electric flux density as $\vec{D}$ [F/m] and the vector of electric field intensity as $\vec{E}$ [V/m or N/C], the total electric flux [C] through a closed surface is found to be equal to the total force charge enclosed by the surface. That is, one finds

$$= \circ \vec{D} \times d\vec{s} = Q_s, \ \vec{D} = e\vec{E},$$

s

where $d\vec{s}$ is the vector surface area, $d\vec{s} = ds\vec{a}_n$, $\vec{a}_n$ is the unit vector which is normal to the surface; e is the permittivity of the medium; Q_s is the total charge enclosed by the surface.

Ohm's law relates the volume charge density $\vec{J}$ and electric field intensity $\vec{E}$; in particular,

$$\vec{J} = s\vec{E},$$

where s is the conductivity [A/V-m], for copper $s = 5.8 \quad 10^7$, and for aluminum $s = 3.5 \quad 10^7$.

The current *i* is proportional to the potential difference, and the resistivity r of the conductor is the ratio between the electric field $\vec{E}$ and the current density $\vec{J}$. Thus,

$$r = \frac{\vec{E}}{\vec{J}}.$$

The resistance *r* of the conductor is related to the resistivity and conductivity by the following formulas

$$r = \frac{rl}{A} \text{ and } r = \frac{l}{sA},$$

where *l* is the length; *A* is the cross-sectional area.

It is important to emphasize that the parameters of NEMS and MEMS vary. Let us illustrate this using the simplest nano-structure used in NEMS and MEMS. In particular, the molecular wire. The resistances of the ware vary due to heating. The resistivity depends on temperature *T* [°C], and

$$r(T) = r_0\left[1 + a_{r1}\left(T \quad T_0\right) + a_{r2}\left(T \quad T_0\right)^2 + ...\right],$$

where a_{r1} and a_{r2} are the coefficients.

As an example, over the small temperature range (up to 160°C) for copper (the wire is filled with copper) at T_0 = 20°C, we have

$$r(T) = 1.7 \quad 10^{-8}\left[1 + 0.0039\left(T \quad 20\right)\right].$$

To study NEMS and MEMS, the basic principles of electromagnetic theory should be briefly reviewed.

The total magnetic flux through the surface is given by

$$= \vec{B} \times d\vec{s},$$

where $\vec{B}$ is the magnetic flux density.

The Ampere circuital law is

$$\vec{B} \times d\vec{l} = m_0 \quad \vec{J} \times d\vec{s},$$

l s

where m_o is the permeability of free space, $m_o = 4 \quad 10^{-7}$ H/m or T-m/A.

For the filamentary current, Ampere's law connects the magnetic flux with the algebraic sum of the enclosed (linked) currents (*net current*) i_n, and

$$\circ \vec{B} \times d\vec{l} = m_o i_n.$$

l

The time-varying magnetic field produces the electromotive force (*emf*), denoted as $\mathcal{J}$, which induces the current in the closed circuit. Faraday's law

relates the *emf*, which is merely the induced voltage due to conductor motion in the magnetic field, to the rate of change of the magnetic flux penetrating in the loop. In approaching the analysis of electromechanical energy transformation in NEMS and MEMS, Lenz's law should be used to find the direction of *emf* and the current induced. In particular, the *emf* is in such a direction as to produce a current whose flux, if added to the original flux, would reduce the magnitude of the *emf*. According to Faraday's law, the induced *emf* in a closed-loop circuit is defined in terms of the rate of change of the magnetic flux as

$$\mathcal{E} = \oint_l \vec{E}(t) \times d\vec{l} = \frac{d}{dt} \int_s \vec{B}(t) \times d\vec{s} = N \frac{d\psi}{dt} = \frac{d\psi}{dt},$$

where N is the number of turns; ψ denotes the flux linkages.

This formula represents the Faraday law of induction, and the induced *emf* (*induced voltage*), as given by

$$\mathcal{E} = \frac{d\psi}{dt} = N \frac{d\Phi}{dt},$$

is a particular interest

The current flows in an opposite direction to the flux linkages. The electromotive force (*energy-per-unit-charge quantity*) represents a magnitude of the potential difference V in a circuit carrying a current. One obtains,

$$V = -ir + \mathcal{E} = ir \frac{d\psi}{dt}.$$

The unit for the *emf* is volts.

The Kirchhoff voltage law states that around a closed path in an electric circuit, the algebraic sum of the *emf* is equal to the algebraic sum of the voltage drop across the resistance.

Another formulation is: the algebraic sum of the voltages around any closed path in a circuit is zero.

The Kirchhoff current law states that the algebraic sum of the currents at any node in a circuit is zero.

The magnetomotive force (*mmf*) is the line integral of the time-varying magnetic field intensity $\vec{H}(t)$; that is,

$$mmf = \oint_l \vec{H}(t) \times d\vec{l}.$$

One concludes that the induced *mmf* is the sum of the induced current and the rate of change of the flux penetrating the surface bounded by the contour. To show that, we apply Stoke's theorem to find the integral form of Ampere's law (second Maxwell's equation), as given by

$$\oint_l \vec{H}(t) \times d\vec{l} = \int_s \vec{J}(t) \times d\vec{s} + \int_s \frac{d\vec{D}(t)}{dt} d\vec{s},$$

where $\vec{J}(t)$ is the time-varying current density vector.

The unit for the magnetomotive force is amperes or ampere-turns

The duality of the *emf* and *mmf* can be observed using

$\mathcal{J} = \oint_l \vec{E}(t) \times d\vec{l}$ and $mmf = \oint_l \vec{H}(t) \times d\vec{l}$.

The inductance (the ratio of the total flux linkages to the current which they link, $L = \frac{N}{i}$) and reluctance (the ratio of the *mmf* to the total flux, $= \frac{mmf}{}$) are used to find *emf* and *mmf*.

Using the following equation for the self-inductance $L = \frac{y}{i}$, we have

$$\mathcal{J} = \frac{dy}{dt} = \frac{d(Li)}{dt} = L\frac{di}{dt} \quad i\frac{dL}{dt}.$$

If $L = const$, one obtains

$$\mathcal{J} = L\frac{di}{dt}.$$

That is, the self-inductance is the magnitude of the self-induced *emf* per unit rate of change of current.

Example 2.2.1.

Find the self-inductances of a nano-solenoid with air-core and filled-core ($m = 100 m_o$). The solenoid has 100 turns (N = 100), the length is 20 nm (l=20 nm), and the uniform circular cross-sectional area is $5 \quad 10^{-18}$ m^2 ($A = 5 \quad 10^{-18}$ m^2).

Solution. The magnetic field inside a solenoid is given by $B = \frac{m_0 Ni}{l}$.

By using $\mathcal{J} = N\frac{d}{dt} = L\frac{di}{dt}$ and applying $= BA = \frac{m_0 NiA}{l}$, one obtains

$$L = \frac{m_0 N^2 A}{l}.$$

Then, L = 3.14 10^{-12} H.

If solenoid is filled with a magnetic material, we have

$L = \frac{mN^2 A}{l}$, and L = 3.14 10^{-9} H.

Example 2.2.2.

Derive a formula for the self-inductance of a torroidal solenoid which has a rectangular cross section ($2a \times b$) and mean radius r.

Solution. The magnetic flux through a cross section is found as

$$= \int_{r}^{r+a} B b dr = \int_{r}^{r+a} \frac{mNi}{2pr} b dr = \frac{mNib}{2p} \int_{r}^{r+a} \frac{1}{r} dr = \frac{mNib}{2p} \ln \left(\frac{r+a}{r} \right).$$

Hence, $$L = \frac{N}{i} = \frac{mN^2 b}{2p} \ln \left(\frac{r+a}{r} \right).$$

By studying the electromagnetic torque $\vec{T}$ [N-m] in a current loop, one obtains the following equation

$$\vec{T} = \vec{M} \quad \vec{B},$$

where $\vec{M}$ denotes the magnetic moment.

Let us examine the torque-energy relations in nano- and microscale actuators. Our goal is to study the magnetic field energy. It is known that the energy stored in the capacitor is $\frac{1}{2} CV^2$, while energy stored in the inductor is $\frac{1}{2} Li^2$. Observe that the energy in the capacitor is stored in the electric field between plates, while the energy in the inductor is stored in the magnetic field within the coils.

Let us find the expressions for energies stored in electrostatic and magnetic fields in terms of field quantities. The total potential energy stored in the electrostatic field is found using the potential difference *V*, and we have

$$W_e = \frac{1}{2} \int_{v} r_v V dv \text{ [J]},$$

where r_v is the volume charge density [C/m^3], $r_v = \vec{} \times \vec{D}$, $\vec{}$ is the curl operator.

This expression for W_e is interpreted in the following way. The potential energy should be found using the amount of work which is required to position the charge in the electrostatic field. In particular, the work is found as the product of the charge and the potential. Considering the region with a continuous charge distribution ($r_v = const$), each charge is replaced by $r_v dv$, and hence the equation $W_e = \frac{1}{2} \int_{v} r_v V dv$ results.

In the Gauss form, using $r_v = \vec{} \times \vec{D}$ and making use $\vec{E} = \vec{} V$, one obtains the following expression for the energy stored in the electrostatic field

$$W_e = \frac{1}{2} \int_{v} \vec{D} \times \vec{E} dv,$$

and the electrostatic volume energy density is $\frac{1}{2} \vec{D} \times \vec{E}$ [J/m^3].

For a linear isotropic medium $W_e = \frac{1}{2} \int_v e\left|\vec{E}\right|^2 dv = \frac{1}{2} \int_v \frac{1}{e}\left|\vec{D}\right|^2 dv$.

The electric field $\vec{E}(x,y,z)$ is found using the scalar electrostatic potential function $V(x,y,z)$ as

$\vec{E}(x,y,z) = -\vec{\nabla} V(x,y,z)$.

In the cylindrical and spherical coordinate systems, we have

$\vec{E}(r,f,z) = -\vec{\nabla} V(r,f,z)$ and $\vec{E}(r,q,f) = -\vec{\nabla} V(r,q,f)$.

Using $W_e = \frac{1}{2} \int_v r_v V dv$, the potential energy which is stored in the electric field between two surfaces (for example, in capacitor) is found to be

$W_e = \frac{1}{2} QV = \frac{1}{2} CV^2$.

Using the principle of virtual work, for the lossless conservative system, the differential change of the electrostatic energy dW_e is equal to the differential change of mechanical energy dW_{mec}; that is

$dW_e = dW_{mec}$.

For translational motion

$dW_{mec} = \vec{F}_e \times d\vec{l}$,

where $d\vec{l}$ is the differential displacement.

One obtains $dW_e = -\vec{\nabla} W_e \times d\vec{l}$.

Hence, the force is the gradient of the stored electrostatic energy, $\vec{F}_e = -\vec{\nabla} W_e$.

In the Cartesian coordinates, we have

$F_{ex} = \frac{\partial W_e}{\partial x}$, $F_{ey} = \frac{\partial W_e}{\partial y}$ and $F_{ez} = \frac{\partial W_e}{\partial z}$.

Example 2.2.3.

Consider the capacitor (the plates have area *A* and they are separated by *x*), which is charged to a voltage *V*. The permittivity of the dielectric is e. Find the stored electrostatic energy and the force F_{ex} in the *x* direction.

Solution. Neglecting the fringing effect at the edges, one concludes that the electric field is uniform, and $E = \frac{V}{x}$. Therefore, we have

$$W_e = \frac{1}{2} \int_v e\left|\vec{E}\right|^2 dv = \frac{1}{2} \int_v e \left(\frac{V}{x}\right)^2 dv = \frac{1}{2} e \frac{V^2}{x^2} Ax = \frac{1}{2} e \frac{A}{x} V^2 = \frac{1}{2} C(x) V^2 .$$

Thus, the force is

$$F_{ex} = \frac{\P W_e}{\P x} = \frac{\P\left(\frac{1}{2}C(x)V^2\right)}{\P x} = \frac{1}{2}V^2\frac{\P C(x)}{\P x}$$

To find the stored energy in the magnetostatic field in terms of field quantities, the following formula is used

$$W_m = \frac{1}{2}\int_v \vec{B}\times\vec{H}dv\,.$$

The magnetic volume energy density is $\frac{1}{2}\vec{B}\times\vec{H}$ [J/m^3].

Using $\vec{B} = m\vec{H}$, one obtains two alternative formulas

$$W_m = \frac{1}{2}\int_v m\left|\vec{H}\right|^2 dv = \frac{1}{2}\int_v \frac{\left|\vec{B}\right|^2}{m}dv\,.$$

To show how the energy concept studied is applied to electromechanical devices, we find the energy stored in inductors. To approach this problem, we substitute $\vec{B} = \vec{\quad} \quad \vec{A}$, and using the following vector identity $\vec{H}\times\vec{\quad} \quad \vec{A} = \vec{\quad}\times\left(\vec{A} \quad \vec{H}\right) + \vec{A}\times\vec{\quad} \quad \vec{H}$, one obtains

$$W_m = \frac{1}{2}\int_v \vec{B}\times\vec{H}dv = \frac{1}{2}\int_v \vec{\quad}\times\left(\vec{A} \quad \vec{H}\right)dv + \frac{1}{2}\int_v \vec{A}\times\vec{\quad} \quad \vec{H}dv$$

$$= \frac{1}{2}\int_s \left(\vec{A} \quad \vec{H}\right)\times d\vec{s} + \frac{1}{2}\int_v \vec{A}\times\vec{J}dv = \frac{1}{2}\int_v \vec{A}\times\vec{J}dv.$$

Using the general expression for the vector magnetic potential $\vec{A}\left(\vec{r}\right)$ [Wb/m], as given by

$$\vec{A}\left(\vec{r}\right) = \frac{m_0}{4p}\int_{v_A}\frac{\vec{J}\left(\vec{r}_A\right)}{x}dv_J\,,\quad \vec{\quad}\times\vec{A} = 0\,,$$

we have

$$W_m = \frac{m}{8p}\int_v\int_{v_J}\frac{\vec{J}\left(\vec{r}_A\right)\times\vec{J}\left(\vec{r}\right)}{x}dv_J dv\,.$$

Here, v_J is the volume of the medium where $\vec{J}$ exists.

The general formula for the self-inductance $i = j$ and the mutual inductance $i \quad j$ of loops *i* and *j* is

$$L_{ij} = \frac{N_i \quad ij}{i_j} = \frac{y_{ij}}{i_j}\,,$$

where y_{ij} is the flux linkage through *i*th coil due to the current in *j*th coil; i_j is the current in *j*th coil.

The *Neumann* formula is applied to find the mutual inductance. We have,

$$L_{ij} = L_{ji} = \frac{m}{4p} \circ\circ_{l_i\ l_j} \frac{d\vec{l}_j \times d\vec{l}_i}{x_{ij}}, i \quad j .$$

Then, using $W_m = \frac{m}{8p} \int_{v\ v_J} \frac{\vec{J}(\vec{r}_A) \times \vec{J}(\vec{r})}{x} dv_J dv$, one obtains

$$W_m = \frac{m}{8p} \int_{l_i\ l_j} \frac{i_j d\vec{l}_j \times i_i d\vec{l}_i}{x_{ij}} .$$

Hence, the energy stored in the magnetic field is found to be

$$W_m = \tfrac{1}{2} i_i L_{ij} i_j .$$

As an example, the energy, stored in the inductor is $W_m = \frac{1}{2} Li^2$.

The differential change in the stored magnetic energy should be found. Using

$$\frac{dW_m}{dt} = \tfrac{1}{2} \left(L_{ij} i_j \frac{di_i}{dt} + L_{ij} i_i \frac{di_j}{dt} + i_i i_j \frac{dL_{ij}}{dt} \right),$$

we have $dW_m = \tfrac{1}{2} \left(L_{ij} i_j di_i + L_{ij} i_i di_j + i_i i_j dL_{ij} \right)$.

For translational motion, the differential change in the mechanical energy is expressed by

$$dW_{mec} = \vec{F}_m \times d\vec{l} .$$

Assuming that the system is conservative (for lossless systems $dW_{mec} = dW_m$), in the rectangular coordinate system we obtain the following equation

$$dW_m = \frac{\partial W_m}{\partial x} dx + \frac{\partial W_m}{\partial y} dy + \frac{\partial W_m}{\partial z} dz = \vec{\nabla} W_m \times d\vec{l} .$$

Hence, the force is the gradient of the stored magnetic energy, and

$$\vec{F}_m = \vec{\nabla} W_m .$$

In the XYZ coordinate system for the translational motion, we have

$$F_{mx} = \frac{\partial W_m}{\partial x}, F_{my} = \frac{\partial W_m}{\partial y} \text{ and } F_{mz} = \frac{\partial W_m}{\partial z} .$$

For the rotational motion, the torque should be used. Using the differential change in the mechanical energy as a function of the angular displacement q, the following formula results if the rigid body (nano- or microactuator) is constrained to rotate about the *z*-axis

$$dW_{mec} = T_e dq ,$$

where T_e is the *z*-component of the electromagnetic torque.

Assuming that the system is lossless, one obtains the following expression for the electromagnetic torque

$$T_e = \frac{\P W_m}{\P q}.$$

Example 2.2.4.

Calculate the magnetic energy of the torroidal microsolenoid if the self-inductance is 1 10 ^{10} H (L=2 10 ^{10} H) when the current is 0.001 A (i=0.001 A).

Solution. The stored field energy is $W_m = \frac{1}{2} Li^2$,

therefore $W_m = \frac{1}{2} 2 \quad 10^{\ 10} \quad 0.001 = 1 \quad 10^{\ 13}$ J.

Example 2.2.5.

Calculate the force developed by the microelectromagnet with the cross-sectional area A if the current $i_a(t)$ in and N coils produces the constant flux $_m$, see Figure 2.2.1.

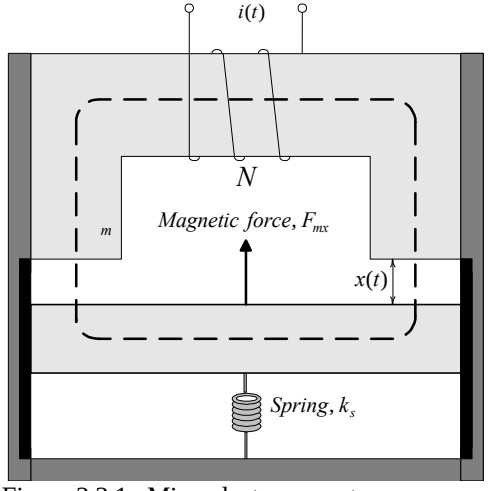


Figure 2.2.1. Microelectromagnet

Solution.

From $W_m = \frac{1}{2} \int_v m \left|\vec{H}\right|^2 dv = \frac{1}{2} \int_v \frac{\left|\vec{B}\right|^2}{m} dv$, for the virtual displacement dy, assuming that the flux is constant and taking into the account the fact that the displacement changes only the magnetic energy stored in the air gaps, we have

$$dW_m = dW_{m\,air\,gap} = 2\frac{B^2}{2\mu_0} A dy = \frac{\mu_m^2}{\mu_0 A} dy .$$

Thus, if μ_m=const, one concludes that the increase of the air gap (dy) leads to increase of the stored magnetic energy, and from $F_{mx} = \frac{\partial W_m}{\partial x}$ one finds the expression for the force

$$\vec{F}_{mx} = \vec{a}_y \frac{\mu_m^2}{\mu_0 A} .$$

The result indicates that the force tends to reduce the air-gap length, and the movable member is attached to the spring which develops the force which opposite to the electromagnetic force.

In nano- and microscale electromechanical motion devices, the coupling (magnetic interaction) between windings that are carrying currents is represented by their mutual inductances. In fact, the current in each winding causes the magnetic field in other windings. The mutually induced *emf* is characterized by the mutual inductance which is a function of the position x or the angular displacement θ. By applying the expression for the coenergy $W_c[i, L(x)]$ or $W_c[i, L(\theta)]$, the developed electromagnetic torque can be easily found. In particular,

$$T_e(i,x) = \frac{\partial W_c[i, L(x)]}{\partial x} \text{ and } T_e(i,x) = \frac{\partial W_c[i, L(\theta)]}{\partial \theta} .$$

Example 2.2.6.

Consider the microelectromagnet which has N turns, see Figure 2.2.2. The distance between the stationary and movable members is denoted as $x(t)$. The mean lengths of the stationary and movable members are l_1 and l_2, and the cross-sectional area is A. Neglecting the leakage flux, find the force exerted on the movable member if the time-varying current $i_a(t)$ is supplied. The permeabilities of stationary and movable members are μ_1 and μ_2.

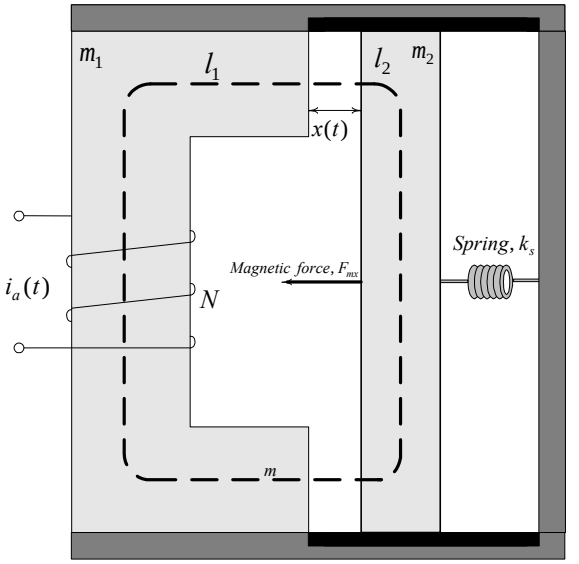


Figure 2.2.2. Schematic of an electromagnet

Solution.

The magnetostatic force is

$$F_{mx} = \frac{\P W_m}{\P x},$$

where $W_m = \frac{1}{2} L i_a^2(t)$.

The magnetizing inductance should be calculated, and we have

$$L = \frac{N}{i_a(t)} = \frac{y}{i_a(t)},$$

where the magnetic flux is $= \dfrac{N i_a(t)}{_1 + \ _x + \ _x + \ _2}$.

The reluctances of the ferromagnetic materials of stationary and movable members $_1$ and $_2$, as well as the reluctance of the air gap $_x$, are found as

$$_1 = \frac{l_1}{m_0 m_1 A}, \quad _2 = \frac{l_2}{m_0 m_2 A} \text{ and } \quad _x = \frac{x(t)}{m_0 A}$$

and the circuit analog with the reluctances of the various paths is illustrated in Figure 2.2.3.

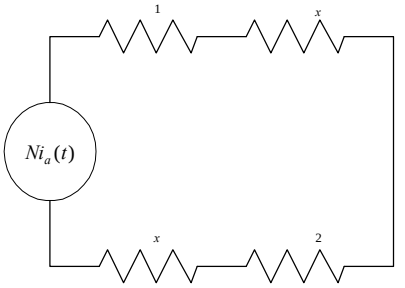


Figure 2.2.3. Circuit analog

By making use the reluctances in the movable and stationary members and air gap, one obtains the following formula for the flux linkages

$$y = N = \frac{N^2 i_a(t)}{\frac{l_1}{m_0 m_1 A} + \frac{2x(t)}{m_0 A} + \frac{l_2}{m_0 m_2 A}},$$

and the magnetizing inductance is a nonlinear function of the displacement. We have

$$L(x) = \frac{N^2}{\frac{l_1}{m_0 m_1 A} + \frac{2x(t)}{m_0 A} + \frac{l_2}{m_0 m_2 A}} = \frac{N^2 m_0 m_1 m_2 A}{m_2 l_1 + m_1 m_2 2x(t) + m_1 l_2}.$$

Using $F_{mx} = \frac{\P W_m}{\P x} = \frac{1}{2}\frac{\P\left(L\left(x(t)\right) i_a^2(t)\right)}{\P x}$, the force in the x direction is found to be

$$F_{mx} = \frac{N^2 m_0 m_1^2 m_2^2 A i_a^2}{m_2 l_1 + m_1 m_2 2x(t) + m_1 l_2}.$$

It should be emphasized that as differential equations must be developed to model the microelectromagnet studied. Using Newton's second law of motion, one obtains

$$\frac{dx}{dt} = v,$$

$$\frac{dv}{dt} = \frac{1}{m} \quad \frac{N^2 m_0 m_1^2 m_2^2 A i_a^2}{m_2 l_1 + m_1 m_2 2x(t) + m_1 l_2} \quad k_s x^2 \div .$$

Example 2.2.7.

Two micro-coils have mutual inductance 0.00005 H (L_{12}=0.00005 H). The current in the first coil is $i_1 = \sqrt{\sin 4t}$. Find the induced *emf* in the second coil.

Solution.

The induced *emf* is given as $\mathcal{E}_2 = L_{12}\frac{di_1}{dt}$.

By using the power rule for the time-varying current in the first coil $i_1 = \sqrt{\sin 4t}$, we have

$$\frac{di_1}{dt} = \frac{2\cos 4t}{\sqrt{\sin 4t}}.$$

Hence, $\mathcal{E}_2 = \frac{0.0001\cos 4t}{\sqrt{\sin 4t}}$.

Basic Foundations in Model Developments of Nano- and Microactuators in Electromagnetic Fields

Electromagnetic theory and mechanics form the basis for the development of NEMS and MEMS models.

The electrostatic and magnetostatic equations in linear isotropic media are found using the vectors of the electric field intensity $\vec{E}$, electric flux density $\vec{D}$, magnetic field intensity $\vec{H}$, and magnetic flux density $\vec{B}$. In addition, one uses the constitutive equations

$\vec{D} = e\vec{E}$ and $\vec{B} = m\vec{H}$

where e is the permittivity; m is the permiability.

The basic equations are given in the Table 1.

Table 2.2.1.

Fundamental Equations of Electrostatic and Magnetostatic Fields

	Electrostatic Model	Magnetostatic Model
Governing equations	$\vec{E}(x,y,z,t) = 0$ $\times\vec{E}(x,y,z,t) = \frac{r_v(x,y,z,t)}{e}$	$\vec{H}(x,y,z,t) = 0$ $\times\vec{H}(x,y,z,t) = 0$
Constitutive equations	$\vec{D} = e\vec{E}$	$\vec{B} = m\vec{H}$

In the static (time-invariant) fields, electric and magnetic field vectors form separate and independent pairs. That is, $\vec{E}$ and $\vec{D}$ are not related to $\vec{H}$ and $\vec{B}$, and vice versa. However, in reality, the electric and magnetic fields are time-varying, and the changes of magnetic field influence the electric field, and vice versa.

The partial differential equations are found by using Maxwell's equations. In particular, four Maxwell's equations in the differential form for time-varying fields are

$$\vec{E}(x,y,z,t) = \mu \frac{\partial \vec{H}(x,y,z,t)}{\partial t},$$

$$\vec{H}(x,y,z,t) = \sigma \vec{E}(x,y,z,t) + \varepsilon \frac{\partial \vec{E}(x,y,z,t)}{\partial t} + \vec{J}(x,y,z,t),$$

$$\nabla \times \vec{E}(x,y,z,t) = \frac{\rho_v(x,y,z,t)}{\varepsilon},$$

$$\nabla \times \vec{H}(x,y,z,t) = 0,$$

where $\vec{E}$ is the electric field intensity, and using the permittivity ε, the electric flux density is $\vec{D} = \varepsilon \vec{E}$; $\vec{H}$ is the magnetic field intensity, and using the permeability μ, the magnetic flux density is $\vec{B} = \mu \vec{H}$; $\vec{J}$ is the current density, and using the conductivity σ, we have $\vec{J} = \sigma \vec{E}$; ρ_v is the volume charge density, and the total electric flux through a closed surface is $= \oint_s \vec{D} \cdot d\vec{s} = \int_v \rho_v dv = Q$ (Gauss's law), while the magnetic flux crossing surface is $= \oint_s \vec{B} \cdot d\vec{s}$.

The *electromotive* and *magnetomotive* forces are found as

$$emf = \oint_l \vec{E} \cdot d\vec{l} = \underbrace{\oint_l (\vec{v} \times \vec{B}) \cdot d\vec{l}}_{\text{motional induction (generation)}} \quad \underbrace{\oint_s \frac{\partial \vec{B}}{\partial t} d\vec{s}}_{\text{transformer induction}} \quad \text{and}$$

$$mmf = \oint_l \vec{H} \cdot d\vec{l} = \oint_s \vec{J} \cdot d\vec{s} + \oint_s \frac{\partial \vec{D}}{\partial t} d\vec{s}.$$

The motional *emf* is a function of the velocity and the magnetic flux density, while the electromotive force induced in a stationary closed circuit is equal to the negative rate of increase of the magnetic flux (transformer induction).

We introduce the vector magnetic potential which is denoted as $\vec{A}$. Using the equation $\vec{B} = \nabla \times \vec{A}$, one finds the following nonhomogeneous vector wave equation

$$\nabla^2 \vec{A} - \mu\varepsilon \frac{\partial^2 \vec{A}}{\partial t^2} = -\mu \vec{J}$$

and the solution gives the waves traveling with the velocity $\frac{1}{\sqrt{\mu\varepsilon}}$.

To develop mathematical models, consider the rotational motion of the bar magnet, current loop, and solenoid in a uniform magnetic field as illustrated in Figure 2.2.4.

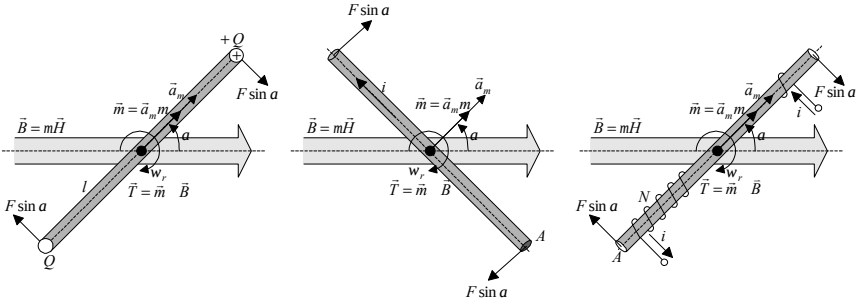


Figure 2.2.4. Clockwise rotation of a magnetic bar, current loop, and solenoid

The torque tends to align the magnetic moment $\vec{m}$ with $\vec{B}$, and $\vec{T} = \vec{m} \quad \vec{B}$.

For a magnetic bar with the length *l*, the pole strength is *Q*.

The magnetic moment is $m = Ql$, and the force is found as $F = QB$.

The electromagnetic torque is found to be

$$T = 2F\,\tfrac{1}{2} l \sin a = QlB \sin a = mB \sin a .$$

Using the vectors, one obtains

$$\vec{T} = \vec{m} \quad \vec{B} = \vec{a}_m m \quad \vec{B} = Ql\vec{a}_m \quad \vec{B}, \tag{2.2.1}$$

where $\vec{a}_m$ is the unit vector in the magnetic moment direction.

For a current loop with the area *A*, the torque is found as

$$\vec{T} = \vec{m} \quad \vec{B} = \vec{a}_m m \quad \vec{B} = iA\vec{a}_m \quad \vec{B}. \tag{2.2.2}$$

For a solenoid with *N* turns, one obtains

$$\vec{T} = \vec{m} \quad \vec{B} = \vec{a}_m m \quad \vec{B} = iAN\vec{a}_m \quad \vec{B}. \tag{2.2.3}$$

The straightforward application of Newton's second law for the rotational motion gives

$$\vec{T} = J\frac{dw_r}{dt}, \tag{2.2.4}$$

where $\vec{T}$ is the net torque; w_r is the angular velocity; *J* is the equivalent moment of inertia\.

The transient evolution of the angular displacement q_r is modeled as

$$\frac{dq_r}{dt} = w_r . \tag{2.2.5}$$

Augmenting equations (2.2.1), (2.2.2) or (2.2.3) with (2.2.4) and (2.2.5), the mathematical model of nano and micro rotational actuators results.

The energy is stored in the magnetic field, and media are classified as diamagnetic, paramagnetic, ferromagnetic, antiferromagnetic, and superparamagnetic. Using the magnetic susceptibility c_m, the magnetization is expressed as

$$\vec{M} = c_m \vec{H} .$$

Magnetization curves should be studied, and the permeability is $m = \frac{B}{H}$. The magnetic field density *B* lags behind the magnetic flux intensity *H*, and this phenomenon is called hysterisis. Thus, the *B*-*H* magnetization curves must be studied.

The per-unit volume magnetic energy stored is $\int_0^B H dB$. The density of the energy stored in the magnetic field is $\frac{1}{2}\vec{B}\times\vec{H}$. If *B* is linearly related to *H*, we have the expression for the total energy stored in the magnetic field as $\frac{1}{2}\int_v \vec{B}\times\vec{H} dv$.

For translational motion, Newton's second law states that the net force acting on the object is related to its acceleration as $\vec{F} = m\vec{a}$. In the XYZ coordinate system, one obtains

$$F_x = ma_x , \qquad F_y = ma_y \text{ and } \qquad F_z = ma_z .$$

The force is the gradient of the stored magnetic energy; that is,

$$\vec{F}_m = \vec{\nabla} W_m .$$

Hence, in the *xyz* directions, we have

$$F_{mx} = \frac{\partial W_m}{\partial x}, F_{my} = \frac{\partial W_m}{\partial y} \text{ and } F_{mz} = \frac{\partial W_m}{\partial z} ,$$

where the stored magnetic energy is found using the volume current density $\vec{J}$

$$W_m = \frac{m}{8p}\int_v \int_{v_A} \frac{\vec{J}(\vec{r}_A)\times\vec{J}(\vec{r})}{R} dv_A dv .$$

Applying the field quantities, we have

$$W_m = \frac{1}{2}\int_v \vec{A}\times\vec{J} dv = \frac{1}{2}\int_v \vec{B}\times\vec{H} dv$$

The magnetic energy density is

$w_m = \frac{1}{2}\vec{A}\times\vec{J} = \frac{1}{2}\vec{B}\times\vec{H}$.

Using Newton's second law and the stored magnetic energy, we have nine highly coupled nonlinear differential equations for the *xyz* translational motion of microactuator. In particular,

$$\frac{dF_{xyz}}{dt} = f_F\left(F_{xyz}, v_{xyz}, x_{xyz}, H\right),$$
$$\frac{dv_{xyz}}{dt} = f_v\left(F_{xyz}, v_{xyz}, x_{xyz}, F_{Lxyz}\right),$$
$$\frac{dx_{xyz}}{dt} = f_x\left(v_{xyz}, x_{xyz}\right), \qquad (2.2.6)$$

where F_{xyz} are the forces developed; v_{xyz} and x_{xyz} are the linear velocities and positions; F_{Lxyz} are the load forces.

The expressions for energies stored in electrostatic and magnetic fields in terms of field quantities should be derived. The total potential energy stored in the electrostatic field is obtained using the potential difference V as $W_e = \frac{1}{2}\int_v r_v V dv$, where the volume charge density is found as $r_v = \vec{\nabla}\times\vec{D}$, $\vec{\nabla}$ is the curl operator.

In the Gauss form, using $r_v = \vec{\nabla}\times\vec{D}$ and making use of $\vec{E} = \vec{\nabla}V$, one obtains the following expression for the energy stored in the electrostatic field $W_e = \frac{1}{2}\int_v \vec{D}\times\vec{E}dv$, and the electrostatic volume energy density is $\frac{1}{2}\vec{D}\times\vec{E}$. For a linear isotropic medium, one finds $W_e = \frac{1}{2}\int_v e\left|\vec{E}\right|^2 dv = \frac{1}{2}\int_v \frac{1}{e}\left|\vec{D}\right|^2 dv$.

The electric field $\vec{E}(x,y,z)$ is found using the scalar electrostatic potential function $V(x,y,z)$ as

$\vec{E}(x,y,z) = \vec{\nabla}V(x,y,z)$.

In the cylindrical and spherical coordinate systems, we have

$\vec{E}(r,f,z) = \vec{\nabla}V(r,f,z)$ and $\vec{E}(r,q,f) = \vec{\nabla}V(r,q,f)$.

Using the principle of virtual work, for the lossless conservative nano- and microelectromechanical systems, the differential change of the electrostatic energy dW_e is equal to the differential change of mechanical

energy dW_{mec}, $dW_e = dW_{mec}$. For translational motion $dW_{mec} = \vec{F}_e \times d\vec{l}$, where $d\vec{l}$ is the differential displacement.

One obtains $dW_e = \vec{\nabla} W_e \times d\vec{l}$.

Hence, the force is the gradient of the stored electrostatic energy,

$\vec{F}_e = \vec{\nabla} W_e$.

In the Cartesian coordinates, we have

$$F_{ex} = \frac{\partial W_e}{\partial x}, \; F_{ey} = \frac{\partial W_e}{\partial y} \text{ and } F_{ez} = \frac{\partial W_e}{\partial z}.$$

Energy conversion takes place in nano- and microscale electromechanical motion devices (actuators and sensors, smart structures), antennas and ICs. We study electromechanical motion devices that convert electrical energy (more precisely electromagnetic energy) to mechanical energy and vise versa (conversion of mechanical energy to electromagnetic energy). Fundamental principles of energy conversion, applicable to nano and micro electromechanical motion devices were studied to provide basic foundations. Using the principle of conservation of energy we can formulate: *for a lossless nano- and microelectromechanical motion devices (in the conservative system no energy is lost through friction, heat, or other irreversible energy conversion) the sum of the instantaneous kinetic and potential energies of the system remains constant.*

The energy conversion is represented in Figure 2.2.5.

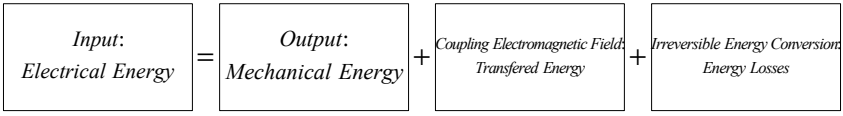


Figure 2.2.5. Energy transfer in nano and micro electromechanical systems

The total energy stored in the magnetic field is found as

$$W_m = \frac{1}{2} \int_v \vec{B} \times \vec{H} dv,$$

where $\vec{B}$ and $\vec{H}$ are related using the permeability μ, $\vec{B} = \mu \vec{H}$.

The material becomes magnetized in response to the external field $\vec{H}$, and the dimensionless magnetic susceptibility χ_m or relative permeability μ_r are used. We have,

$$\vec{B} = \mu \vec{H} = \mu_0 (1 + \chi_m) \vec{H} = \mu_0 \mu_r \vec{H} = \mu \vec{H}.$$

Based upon the value of the magnetic susceptibility χ_m, the materials are classified as

diamagnetic, $c_m \quad 1 \quad 10^{5}$ ($c_m = \quad 9.5 \quad 10^{6}$ for copper, $c_m = \quad 3.2 \quad 10^{5}$ for gold, and $c_m = \quad 2.6 \quad 10^{5}$ for silver);

paramagnetic, $c_m \quad 1 \quad 10^{4}$ ($c_m = 1.4 \quad 10^{3}$ for Fe_2O_3, and $c_m = 1.7 \quad 10^{3}$ for Cr_2O_3);

ferromagnetic, $|c_m| >> 1$ (iron, nickel and cobalt, Neodymium Iron Boron and Samarium Cobalt permanent magnets) .

The magnetization behavior of the ferromagnetic materials is mapped by the magnetization curve, where *H* is the externally applied magnetic field, and *B* is total magnetic flux density in the medium. Typical *B*-*H* curves for hard and soft ferromagnetic materials are given in Figure 2.2.6, respectively.

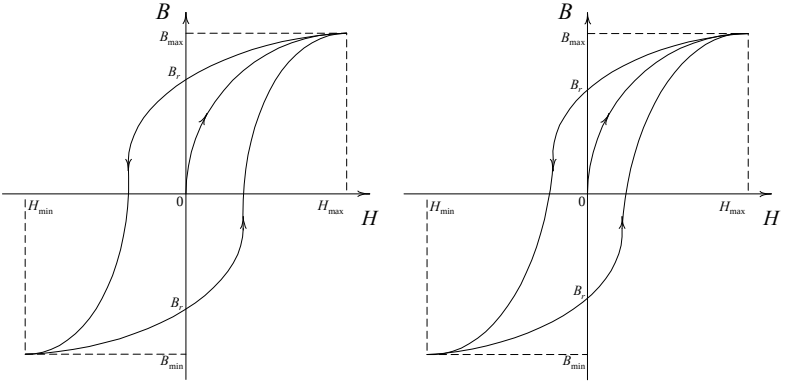


Figure 2.2.6. *B*-*H* curves for hard and soft ferromagnetic materials

The *B* versus *H* curve allows one to establish the energy analysis. Assume that initially $B_0 = 0$ and $H_0 = 0$. Let *H* increases form $H_0 = 0$ to $H_{\max}$. Then, *B* increases from $B_0 = 0$ until the maximum value of *B*, denoted as $B_{\max}$, is reached. If then *H* decreases to $H_{\min}$, *B* decreases to $B_{\min}$ through the remanent value B_r (the so-called the residual magnetic flux density) along the different curve, see Figure 2.18. For variations of *H*, $H \quad [H_{\min} \; H_{\max}]$, *B* changes within the *hysteresis loop*, and $B \quad [B_{\min} \; B_{\max}]$.

In the per-unit volume, the applied field energy is $W_F = \circ_B HdB$, while the stored energy is expressed as $W_c = \circ_H BdH$.

In the volume *v*, we have the following expressions for the field and stored energy

$W_F = v\circ_B HdB$ and $W_c = v\circ_H BdH$.

A complete *B* versus *H* loop should be considered, and the equations for field and stored energy represent the areas enclosed by the corresponding curve. It should be emphasized that each point of the *B* versus *H* curve represent the total energy.

In ferromagnetic materials, time-varying magnetic flux produces core losses which consist of hysteresis losses (due to the hysteresis loop of the *B*-*H* curve) and the eddy-current losses, which are proportional to the current frequency and lamination thickness. The area of the hysteresis loop is related to the hysteresis losses. Soft ferromagnetic materials have narrow hysteresis loop and they are easily magnetized and demagnetized. Therefore, the lower hysteresis losses, compared with hard ferromagnetic materials, result.

For electromechanical motion devices, the flux linkages are plotted versus the current because the current and flux linkages are used rather than the flux intensity and flux density. In nano- and microectromechanical motion devices almost all energy is stored in the air gap. Using the fact that the air is a conservative medium, one concludes that the coupling filed is lossless.

Figure 2.2.7 illustrates the nonlinear magnetizing characteristic (normal magnetization curve), and the energy stored in the magnetic field is $W_F = \circ_y idy$, while the coenergy is found as $W_c = \circ_i ydi$. The total energy is

$$W_F + W_c = \circ_y idy + \circ_i ydi = yi.$$

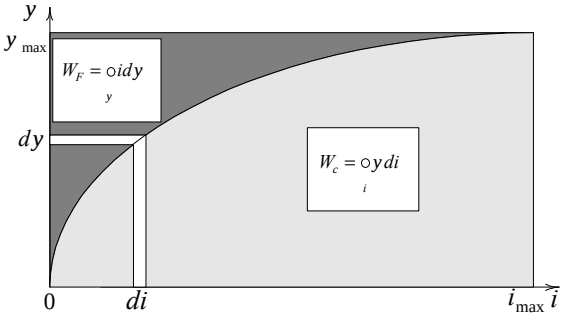


Figure 2.2.7. Magnetization curve and energies

The flux linkages is the function of the current i and position x (for translational motion) or angular displacement q (for rotational motion). That is, $y = f(i,x)$ or $y = f(i,q)$. The current can be found as the nonlinear function of the flux linkages and position or angular displacement. Hence,

$$dy = \frac{\partial y(i,x)}{\partial i}di + \frac{\partial y(i,x)}{\partial x}dx, \quad dy = \frac{\partial y(i,q)}{\partial i}di + \frac{\partial y(i,q)}{\partial q}dq,$$

and $$di = \frac{\partial i(y,x)}{\partial y}dy + \frac{\partial i(y,x)}{\partial x}dx, \quad di = \frac{\partial i(y,q)}{\partial y}dy + \frac{\partial i(y,q)}{\partial q}dq.$$

Therefore,

$$W_F = \int_y i\,dy = \int_i i\frac{\partial y(i,x)}{\partial i}di + \int_x i\frac{\partial y(i,x)}{\partial x}dx,$$

$$W_F = \int_y i\,dy = \int_i i\frac{\partial y(i,q)}{\partial i}di + \int_q i\frac{\partial y(i,q)}{\partial q}dq,$$

and $$W_c = \int_i y\,di = \int_y y\frac{\partial i(y,x)}{\partial y}dy + \int_x y\frac{\partial i(y,x)}{\partial x}dx,$$

$$W_c = \int_i y\,di = \int_y y\frac{\partial i(y,q)}{\partial y}dy + \int_q y\frac{\partial i(y,q)}{\partial q}dq.$$

Assuming that the coupling field is lossless, the differential change in the mechanical energy (which is found using the differential displacement $d\vec{l}$ as $dW_{mec} = \vec{F}_m \times d\vec{l}$) is related to the differential change of the coenergy. For displacement dx at constant current, one obtains $dW_{mec} = dW_c$, and hence, the electromagnetic force is $F_e(i,x) = \frac{\partial W_c(i,x)}{\partial x}$.

For rotational motion, the electromagnetic torque is

$$T_e(i,q) = \frac{\partial W_c(i,q)}{\partial q}.$$

Micro- and meso-scale structures, as well as thin magnetic films, exhibit anisotropy. Consider the anisotropic ferromagnetic element in the Cartesian (rectangular) coordinate systems as shown in Figure 2.2.8.

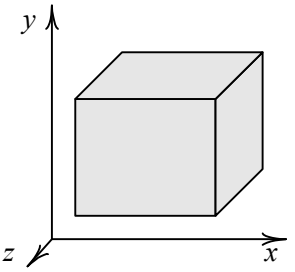


Figure 2.2.8. Material in the *xyz* coordinate system

The permeability is $m(x,y,z) = \begin{matrix} m_{xx} & m_{xy} & m_{xz} \\ m_{yx} & m_{yy} & m_{yz} \\ m_{zx} & m_{zy} & m_{zz} \end{matrix}$, and therefore,

$$\vec{B} = m(x,y,z)\vec{H}, \quad \begin{matrix} B_x \\ B_y \\ B_z \end{matrix} = \begin{matrix} m_{xx} & m_{xy} & m_{xz} \\ m_{yx} & m_{yy} & m_{yz} \\ m_{zx} & m_{zy} & m_{zz} \end{matrix} \begin{matrix} H_x \\ H_y \\ H_z \end{matrix}.$$

The analysis of anisotropic nano- and microscale actuators and sensors can be performed. Some actuators and sensors can be studied assuming that the media is linear, homogeneous, and isotropic. Unfortunately, this assumption is not valid in general.

Control of microactuators position and linear velocity, angular displacement and angular velocity, is established by changing *H*. In (2.2.6), the magnetic field intensity can be considered as a control. However, the electromagnetic field is developed by ICs or antennas. Hence, the microICs or microantenna dynamics have to be integrated in (2.2.6). Thus, microscale antennas and ICs must be thoroughly considered.

Consider the microactuator controlled by the microantenna. Assume that the linear isotropic media has permittivity $e_0 e_m$ and permeability $m_0 m_m$.

The force is calculated using the stress energy tensor $\vec{\vec{T}}_{ab}$ which is given in terms of the electromagnetic field as

$$\vec{\vec{T}}_{ab} = e_0 e_m E_a E_b + m_0 m_m H_a H_b \quad \tfrac{1}{2} d_{ab} \left(e_0 e_m E_g E_g + m_0 m_m H_g H_g \right),$$

where d_{ab} is the Kronecker delta-function, defined as $d_{ab} = \begin{matrix} 1 \text{ if } a = b \\ 0 \text{ if } a \quad b \end{matrix}$.

The electromagnetic field tensor is expressed as

$$\vec{\vec{F}}_{ab} = \begin{matrix} 0 & \vec{E}_x & \vec{E}_y & \vec{E}_z \\ \vec{E}_x & 0 & \vec{B}_z & \vec{B}_y \\ \vec{E}_y & \vec{B}_z & 0 & \vec{B}_x \\ \vec{E}_z & \vec{B}_y & \vec{B}_x & 0 \end{matrix},$$

and Maxwell's equation can be expressed in the tensor form.

Then, the electromagnetic force is found as

$$\vec{F} = \int_s \vec{\vec{T}}_{ab} d\vec{s} .$$

The results derived can be viewed using the energy analysis, and one has

$$\vec{F}(\vec{\mathbf{r}}) = \qquad (\vec{\mathbf{r}}), \quad (\vec{\mathbf{r}}) = \frac{e_0 e_m}{2} \int_s \vec{E} \times \vec{E} dv + \frac{1}{2 m_0 m_m} \int_s \vec{H} \times \vec{H} dv .$$

References

1. Hayt W. H., *Engineering Electromagnetics*, McGraw-Hill, New York, 1989.
2. Krause J. D and Fleisch D. A, *Electromagnetics With Applications*, McGraw-Hill, New York, 1999.
3. Krause P. C. and Wasynczuk O., *Electromechanical Motion Devices*, McGraw-Hill, New York, 1989.
4. Lyshevski S. E., *Electromechanical Systems, Electric Machines, and Applied Mechatronics*, CRC Press, FL, 1999.
5. Paul C. R., Whites K. W., and Nasar S. A., *Introduction to Electromagnetic Fields*, McGraw-Hill, New York, 1998.
6. White D. C. and Woodson H. H., *Electromechanical Energy Conversion*, Wiley, New York, 1959.

2.3. CLASSICAL MECHANICS AND ITS APPLICATION

With advanced molecular computer-aided-design tools, one can design, analyze, and evaluate three-dimensional (3-D) nanostructures in the steady-state. However, the comprehensive analysis in the time domain needs to be performed. That is, the designer must study the dynamic evolution of NEMS and MEMS. Conventional methods of molecular mechanics do not allow one to perform numerical analysis of complex NEMS and MEMS in time-domain, and even 3-D modeling is restricted to simple structures. Our goal is to develop a fundamental understanding of electromechanical and electromagnetic processes in nano- and microscale structures. An addition, the basic theoretical foundations will be developed and used in analysis of NEMS and MEMS from systems standpoints. That is, we depart from the subsystem analysis and study NEMS and MEMS as dynamics systems.

From modeling, simulation, analysis, and visualization standpoints, NEMS and MEMS are very complex. In fact, NEMS and MEMS are modeled using advanced concepts of quantum mechanics, electromagnetic theory, structural dynamics, thermodynamics, thermochemistry, etc. It was illustrated that NEMS and MEMS integrate a great number of components (subsystems), and mathematical model development is an extremely challenging problem because the commonly used conventional methods, assumptions, and simplifications may not be applied to NEMS and MEMS (for example, the Newtonian mechanics are not applicable to the molecular-scale analysis, and Maxwell's equations must be used to study the electromagnetic phenomena). As the result, partial differential equations describe large-scale multivariable mathematical models of MEMS and NEMS. The visualization issues must be addressed to study the complex tensor data (tensor field). Techniques and software for visualizing scalar and vector field data are available to visualize the data in three dimensions. In contrast, techniques to visualize tensor fields are not available due to the complex, multivariate nature of the data, and the fact that no commonly used experimental analogy exists for visualizing tensor data. The second-order tensor fields consist of $3 \; 3$ matrices defined at each node in a computational grid. Tensor field variables can include stress, viscous stress, rate of strain, and momentum (tensor variables in conventional structural dynamics include stress and strain). The tensor field can be simplified and visualized as a scalar field. Alternatively, the individual vectors that comprise the tensor field can be analyzed. However, these simplifications result in the loss of valuable information needed to analyze complex tensor fields. Vector fields can be visualized using streamlines that depict a subset of the data. Hyperstreamlines, as an extension of the streamlines to the second-order tensor fields, provide one with a continuous representation of the tensor field along a three-dimensional path. Due to obvious limitations and scope, this book does not cover the tensor field topologies, and through this brief

discussion of the resultant visualization, the author emphasizes the multidisciplinary nature and complexity of the phenomena in NEMS and MEMS.

While some results have been thoroughly studied, many important aspects have not been approached and researched, primarily due to the multidisciplinary nature and complexity of NEMS and MEMS. The major objectives of this book are to study the fundamental theoretical foundations, develop innovative concepts in structural design and optimization, perform modeling and simulation, as well as solve the motion control problem and validate the results. To develop mathematical models, we augment nano- or microactuator/sensor and circuitry dynamics (the dynamics can be studied at the nano and micro scales). Newtonian and quantum mechanics, Lagrange's and Hamilton's concepts, and other cornerstone theories are used to model NEMS and MEMS dynamics in the time domain. Taking note of these basic principles and laws, nonlinear mathematical models are found to perform comprehensive analysis and design. The control mechanisms and decision making are discussed, and control algorithms must be synthesized to attain the desired specifications and requirements imposed on the performance. It is evident that nano- and microsystem features must be thoroughly considered when approaching modeling, simulation, analysis, and design. The ability to find mathematical models is a key problem in NEMS and MEMS analysis and optimization, synthesis and control, manufacturing, and commercialization. For MEMS, using electromagnetic theory and electromechanics, we develop adequate mathematical models to attain the design objectives. The proposed approach, which augments electromagnetics and electromechanics, allows the designer to solve a much broader spectrum of problems compared with finite-element analysis because an interactive electromagnetic-mechanical-ICs analysis is performed. The developed theoretical results are verified to demonstrate.

In this book the author studies large-scale NEMS and MEMS (actuators and sensors have been primarily studied and analyzed from the fabrication standpoints) and thorough fundamental theory is developed. Applying the theoretical foundations to analyze and regulate in the desired manner the energy or information flows in NEMS and MEMS, the designer is confronted with the need to find adequate mathematical models of the phenomena, and design NEMS and MEMS configurations. Mathematical models can be found using basic physical concepts. In particular, in electrical, mechanical, fluid, or thermal systems, the mechanism of storing, dissipating, transforming, and transferring energies is analyzed. We will use the Lagrange equations of motion, Kirchhoff's and Newton's laws, Maxwell's equations, and quantum theory to illustrate the model developments. It was emphasized that NEMS and MEMS integrate many components and subsystems. One can reduce interconnected systems to simple, idealized subsystems (components). However, this idealization is impractical. For example, one cannot study

nano- and microscale actuators and sensors without studying subsystems (devices) to actuate and control these transducers. That is, NEMS and MEMS integrate mechanical and electromechanical motion devices (actuators and sensors), power converters and antennas, processors and IO devices, etc. One of the primary objectives of this book is to illustrate how one can develop comprehensive mathematical models of NEMS and MEMS using basic principles and laws. Through illustrative examples, differential equations will be found to model dynamic systems.

Based upon the synthesized NEMS and MEMS architectures, to analyze and regulate in the desired manner the energy or information flows, the designer needs to find adequate mathematical models and optimize the performance characteristics through the design of control algorithms. Some mathematical models can be found using basic foundations and mathematical theory to map the dynamics of some processes, and system evolution is not developed yet. In this section we study electrical, mechanical, fluid, and thermal systems, the mechanism of storing, dissipating, transforming, and transferring energies in actuators and sensors which can be manufactured using a large variety of different nano-, micro-, and miniscale technologies. In this section we will use the Lagrange equations of motion, as well as Kirchhoff's and Newton's laws to illustrate the model developments applicable to a large class of nano-, micro-, and miniscale transducers. It has been illustrated that one cannot reduce interconnected systems (NEMS and MEMS) to simple, idealized sub-systems (components). For example, one cannot study actuators and smart structures without studying the mechanism to regulate these actuators, and ICs and antennas must be integrated as well. These ICs and antennas are controlled by the processor, which receives the information from sensors. The primary objective of this chapter is to illustrate how one can develop mathematical models of dynamic systems using basic principles and laws. Through illustrative examples, differential equations will be found and simulated.

Nano- and microelectromechanical systems must be studied using the fundamental laws and basic principles of mechanics and electromagnetics. Let us identify and study these key concepts to illustrate the use of cornerstone principles. The study of the motion of systems with the corresponding analysis of forces that cause motion is our interest.

2.3.1. Newtonian Mechanics

Newtonian Mechanics: Translational Motion

The equations of motion for mechanical systems can be found using Newton's second law of motion. Using the position (displacement) vector $\vec{\mathbf{r}}$, the Newton equation in the vector form is given as

$$\vec{F}(t,\vec{\mathbf{r}}) = m\vec{a}\,, \tag{2.3.1}$$

where $\vec{F}(t,\vec{\mathbf{r}})$ is the vector sum of all forces applied to the body ($\vec{F}$ is called the *net* force); $\vec{a}$ is the vector of acceleration of the body with respect to an inertial reference frame; m is the mass of the body.

From (2.3.1), in the Cartesian system (*xyz* coordinates) we have

$$\vec{F}(t,\vec{\mathbf{r}}) = m\vec{a} = m\frac{d\vec{\mathbf{r}}^2}{dt^2} = m\begin{matrix}\frac{d\vec{x}^2}{dt^2}\\ \frac{d\vec{y}^2}{dt^2}\\ \frac{d\vec{z}^2}{dt^2}\end{matrix}, \quad \begin{matrix}\vec{a}_x\\ \vec{a}_y\\ \vec{a}_z\end{matrix} = \begin{matrix}\frac{d\vec{x}^2}{dt^2}\\ \frac{d\vec{y}^2}{dt^2}\\ \frac{d\vec{z}^2}{dt^2}\end{matrix}.$$

In the Cartesian coordinate system, Newton's second law is expressed as

$F_x = ma_x$, $F_y = ma_y$, and $F_z = ma_z$.

It is worth noting that $m\vec{a}$ represents the magnitude and direction of the applied net force acting on the object. Hence, $m\vec{a}$ is not a force.

A body is at equilibrium (the object is at rest or is moving with constant speed) if $\vec{F} = 0$.

Newton's second law in terms of the linear momentum, which is found as $\vec{p} = m\vec{v}$, is given by

$$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d(m\vec{v})}{dt},$$

where $\vec{v}$ is the vector of the object velocity.

Thus, the force is equal to the rate of change of the momentum. The object or particle moves uniformly if $\frac{d\vec{p}}{dt} = 0$ (thus, $\vec{p} = const$).

Newton's laws are extended to multi-body systems, and the momentum of a system of N particles is the vector sum of the individual momenta. That is,

$$\vec{P} = \sum_{i=1}^{N} \vec{p}_i\,.$$

Consider the multi-body system of N particles. The position (displacement) is represented by the vector **r** which in the Cartesian coordinate system has the components *x*, *y* and *z*. Taking note of the expression for the potential energy $(\vec{\mathbf{r}})$, one has for the conservative mechanical system

$$\vec{F}(\vec{\mathbf{r}}) = \qquad (\vec{\mathbf{r}})\,.$$

Therefore, the work done per unit time is

$$\frac{dW}{dt} = \quad \vec{F}(\vec{\mathbf{r}})\frac{d\vec{\mathbf{r}}}{dt} = \quad (\vec{\mathbf{r}})\frac{d\vec{\mathbf{r}}}{dt} = \quad \frac{d\ (\vec{\mathbf{r}})}{dt}.$$

From Newton's second law one obtains

$$m\vec{a} \qquad \vec{F}(\vec{\mathbf{r}}) = m\frac{d^2\vec{\mathbf{r}}}{dt^2} \qquad \vec{F}(\vec{\mathbf{r}}) = 0,$$

hence, for a conservative system we have

$$m\frac{d^2\vec{\mathbf{r}}}{dt^2} + \quad (\vec{\mathbf{r}}) = 0.$$

For the system of N particles, the equations of motion are

$$m_N\frac{d^2\vec{\mathbf{r}}_N}{dt^2} + \quad (\vec{\mathbf{r}}_N) = 0,$$

or $\quad m_i\frac{d^2(\vec{x}_i, \vec{y}_i, \vec{z}_i)}{dt^2} + \frac{(\vec{x}_i, \vec{y}_i, \vec{z}_i)}{(\vec{x}_i, \vec{y}_i, \vec{z}_i)} = 0, i = 1,..., N.$

The total kinetic energy of the particle is $= \frac{1}{2}mv^2$, and for N particles, one has

$$\frac{d\vec{x}_i}{dt}, \frac{d\vec{y}}{dt}_i, \frac{d\vec{z}_i}{dt} \div = \frac{1}{2}\sum_{i=1}^{N} m_i \ \frac{d\vec{x}_i}{dt}, \frac{d\vec{y}}{dt}_i, \frac{d\vec{z}_i}{dt} \div.$$

Furthermore, we have $m_i\frac{d(\vec{x}_i, \vec{y}_i, \vec{z}_i)}{dt} = \frac{\frac{d\vec{x}_i}{dt}, \frac{d\vec{y}_i}{dt}, \frac{d\vec{z}_i}{dt} \div}{\frac{d\vec{x}_i}{dt}, \frac{d\vec{y}_i}{dt}, \frac{d\vec{z}_i}{dt} \div}.$

Using the generilized coordinates $(q_1,..., q_n)$ and generalized velocities $\frac{dq_1}{dt},..., \frac{dq_n}{dt} \div$, one finds the total kinetic $\quad q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt} \div$ and potential $\quad (q_1,..., q_n)$ energies. Hence, using the expressions for the total kinetic and potential energies, Newton's second law of motion can be given as

$$\frac{d}{dt}\ \frac{\P}{\P\dot{q}_i} \div + \frac{\P}{\P q_i} = 0.$$

That is, the generalized coordinates q_i are used to model multibody systems, and $(q_1,..., q_n) = (\vec{x}_1, \vec{y}_1, \vec{z}_1, ..., \vec{x}_N, \vec{y}_N, \vec{z}_N)$.

The obtained results are connected to the Lagrange equations of motion, which will be studied later.

For rotational motion, the net torque and angular acceleration must are used. The rotational analog of Newton's second law for a rigid bogy is

$$\vec{T} = J\vec{a},$$

where $\vec{T}$ is the *net* torque; J is the moment of inertia (*rotational inertia*); $\vec{a}$ is the angular acceleration vector, $\vec{a} = \frac{d}{dt}\frac{d\vec{q}}{dt} = \frac{d^2\vec{q}}{dt^2} = \frac{d\vec{w}}{dt}$; $\vec{q}$ is the angular displacement; w denotes the angular velocity.

The angular momentum of the system $\vec{L}_M$ is expressed as

$$\vec{L}_M = \vec{r} \quad \vec{p} = \vec{r} \quad m\vec{v},$$

and $$\vec{T} = \frac{d\vec{L}_M}{dt} = \vec{r} \quad \vec{F},$$

where $\vec{r}$ is the position vector with respect to the origin.

For the rigid body, rotating around the axis of symmetry, we have

$$\vec{L}_M = J\vec{w}.$$

Example 2.3.1.

A micro-motor has the equivalent moment of inertia $J = 5 \quad 10^{-20}$ kg-m^2. Let the angular velocity of the rotor is $w_r = 10t^{1/5}$. Find the angular momentum and the developed electromagnetic torque as functions of time. The load and friction torques are zero.

Solution.

The angular momentum is found as $L_M = Jw_r = 5 \quad 10^{-19} t^{1/5}$.

The developed electromagnetic torque is $T_e = \frac{dL_M}{dt} = 1 \quad 10^{-19} t^{-4/5}$.

From Newtonian mechanics one concludes that the applied net force plays a central role in quantitatively describing the motion. An alternative analysis of motion can be performed in terms of the energy or momentum quantities, which are conserved. The principle of conservation of energy states that energy can only be converted from one form to another. Kinetic energy is associated with motion, while potential energy is associated with position. The sum of the kinetic ( ), potential ( ), and dissipated (D) energies is called the total energy of the system ($_T$), which is conserved, and the total amount of energy remains constant; that is,

$$_T = \quad + \quad + D = const.$$

For example, consider the translational motion of a body which is attached to an ideal spring that obeys Hooke's law. Neglecting friction, one obtains the following expression for the total energy

$$_T = \quad + \quad = \tfrac{1}{2}(mv^2 + k_s x^2) = const\ .$$

Here, the translational kinetic energy is $= \tfrac{1}{2}mv^2$; the elastic potential energy is $= \tfrac{1}{2}k_s x^2$; k_s is the force constant of the spring; x is the displacement.

For rotating spring, we have

$$_T = \quad + \quad = \tfrac{1}{2}(Jw^2 + k_s q^2) = const\ ,$$

where the rotational kinetic energy is $= \tfrac{1}{2}Jw^2$ and the elastic potential energy is obtained as $= \tfrac{1}{2}k_s q^2$.

The kinetic energy of a rigid body having translational and rotational components of motion is found to be

$$= \tfrac{1}{2}(mv^2 + Jw^2)\ .$$

That is, motion of the rigid body is represented as a combination of translational motion of the center of mass and rotational motion about the axis through the center of mass. The moment of inertia depends upon how the mass is distributed with respect to the axis, and J is different for different axes of rotation. If the body is uniform in density, J can be easily calculated for regularly shaped bodies in terms of their dimensions. For example, a rigid cylinder with mass m (which is uniformly distributed), radius R, and length l, has the following horizontal and vertical moments of inertia $J_{horizontal} = \tfrac{1}{2}mR^2$ and $J_{vertical} = \tfrac{1}{4}mR^2 + \tfrac{1}{12}ml^2$. The *radius of gyration* can be found for irregularly shaped objects, and the moment of inertia can be easily obtained.

In electromechanical motion devices, the force and torque are thoroughly studied. Assuming that the body is rigid and the moment of inertia is constant, one has

$$\vec{T}d\vec{q} = J\vec{a}d\vec{q} = J\frac{d\vec{w}}{dt}d\vec{q} = J\frac{d\vec{q}}{dt}d\vec{w} = J\vec{w}d\vec{w}\ .$$

The total work, as given by

$$W = \int_{q_0}^{q_f}\vec{T}d\vec{q} = \int_{w_0}^{w_f}J\vec{w}d\vec{w} = \tfrac{1}{2}(Jw_f^2 \quad Jw_0^2)\ ,$$

represents the change of the kinetic energy.

Furthermore,

$$\frac{dW}{dt} = \vec{T}\frac{d\vec{q}}{dt} = \vec{T} \quad \vec{w},$$

and the power is defined by

$$P = \vec{T} \quad \vec{w}.$$

This equation is an analog of $P = \vec{F} \quad \vec{v}$, which is applied for translational motion.

Example 2.3.2.

Consider a micro-positioning table actuated by a micromotor. How much work is required to accelerate a 2 mg payload (m = 2 mg) from v_0 = 0 m/sec to v_f = 1 m/sec?

Solution.

The work needed is calculated as

$$W = \tfrac{1}{2}(mv_f^2 \quad mv_0^2) = \tfrac{1}{2}2 \quad 10^{\ 6} \quad 1^2 = 1 \quad 10^{\ 6} \text{ J.} \qquad \sim$$

Example 2.3.3.

The rated power and angular velocity of a micromotor are 0.001 W and 100 rad/sec. Calculate the rated electromagnetic torque.

Solution.

The electromagnetic torque is

$$T_e = \frac{P}{w_r} = \frac{0.001}{100} = 1 \quad 10^{\ 5} \text{ N-m.} \qquad \sim$$

Example 2.3.4.

Consider a body of mass m in the XY coordinate system. The force $\vec{F}_a$ is applied in the x direction. Neglecting *Coulomb* and static friction, and assuming that the viscous friction force is $F_{fr} = B_v \frac{dx}{dt}$, find the equations of motion. Here B_v is the viscous friction coefficient.

Solution.

The free-body diagram developed is illustrated in Figure 2.3.1.

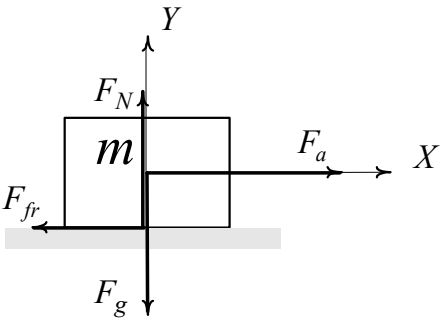


Figure 2.3.1. Free-body diagram

The sum of the forces, acting in the *y* direction, is expressed as

$$\vec{F}_Y = \vec{F}_N \quad \vec{F}_g ,$$

where $\vec{F}_g = mg$ is the gravitational force acting on the mass m; $\vec{F}_N$ is the normal force which is equal and opposite to the gravitational force.

From (2.3.1), the equation of motion in the *y* direction is expressed as

$$\vec{F}_N \quad \vec{F}_g = ma_y = m\frac{d^2y}{dt^2},$$

where a_y is the acceleration in the *y* direction, $a_y = \dfrac{d^2y}{dt^2}$.

Making use $\vec{F}_N = \vec{F}_g$, we have

$$\frac{d^2y}{dt^2} = 0 .$$

The sum of the forces acting in the *x* direction is found using the applied force $\vec{F}_a$ and the friction force $\vec{F}_{fr}$; in particular, we have

$$\vec{F}_X = \vec{F}_a \quad \vec{F}_{fr} .$$

The applied force can be time-invariant $\vec{F}_a = const$ or time-varying $\vec{F}_a(t) = f(t,x,y,z)$. For example,

$$\vec{F}_a(t) = x\sin(6t \quad 4)e^{\ 0.5t} + \frac{dy}{dt}t^2 + z^3\cos\ \frac{dx}{dt}t \quad y^2t^4 \div .$$

Using (2.1), the equation motion in the *x* direction is found to be

$$\vec{F}_a \quad \vec{F}_{fr} = ma_x = m\frac{d^2x}{dt^2},$$

where a_x is the acceleration in the X direction, $a_x = \frac{d^2x}{dt^2}$, and the velocity in the X direction is $v = \frac{dx}{dt}$.

Assuming that the *Coulomb* and static friction can be neglected, the friction force, as a function of the viscous friction coefficient B_v and velocity $v = \frac{dx}{dt}$, is given by $F_{fr} = B_v \frac{dx}{dt}$.

Hence, one obtains the second-order nonlinear differential equation to map the body dynamics in the *x* direction

$$\frac{d^2x}{dt^2} = \frac{1}{m}\left(F_a - B_v \frac{dx}{dt}\right),$$

A set of two first-order linear differential equations results, and

$$\frac{dx}{dt} = v,$$

$$\frac{dv}{dt} = \frac{1}{m}\left(F_a - B_v v\right).$$

The application of Newton's law leads to the partial differential equations. To illustrate this, we consider two examples.

Example 2.3.5.

The elastic membrane is illustrated in Figure 2.3.2. Derive the mathematical model to model the rectangular membrane vibration. That is, the goal is to study the time varying membrane deflection *d*(*t*,*x*,*y*) in the *xy* plane. The mass of the undeflected membrane per unit area r is constant (homogeneous membrane).

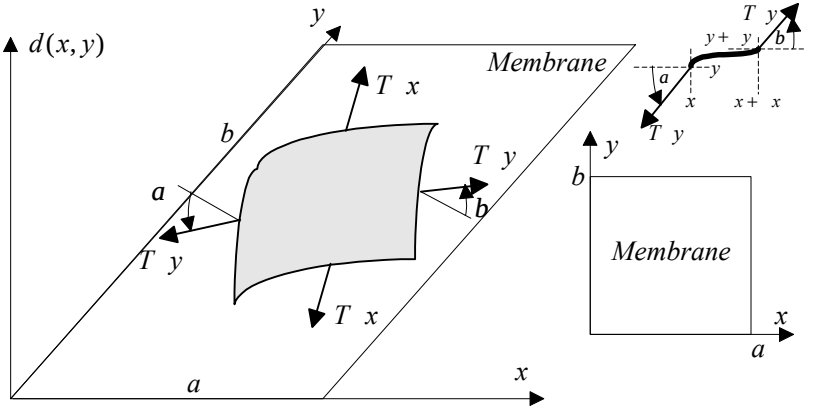


Figure 2.3.2. Vibrating rectangular membrane

Solution.

Assume that the membrane is perfectly flexible. For small deflections, the tension T (the force per unit length) is the same at all points in all directions, and suppose that T is constant during the motion. It should be emphasized that because the deflection of the membrane is small compared with the membrane size ab, the inclination angles are small.

Taking note of these assumptions, the forces acting on the sides are approximated as $F_x = T\ x$ and $F_y = T\ y$. The membrane is assumed to be perfectly flexible, therefore, forces F_x and F_y are tangential to the membrane.

The horizontal components of the forces are found as the cosine functions of the inclination angles. The horizontal components at the opposite sides (right and left) are equal because angles a and b are small. Thus, the membrane motion in the horizontal direction can be neglected.

The vertical components of the forces are $T\ y \sin b$ and $T\ y \sin a$.

Using Newton's second law of motion, the net force must be found. We have the following expression

$$F = T\ y\big(d_x(x+\ x, y_1)\ \ d_x(x, y_2)\big) + T\ x\big(d_y(x_1, y+\ y)\ \ d_y(x_2, y)\big)$$

Thus, two-dimensional partial (wave) differential equation is

$$\frac{^2 d(t,x,y)}{t^2} = \frac{T}{r}\ \frac{^2 d(t,x,y)}{x^2} + \frac{^2 d(t,x,y)}{y^2} \div = \frac{T}{r}\ \ ^2 d(t,x,y).$$

Using initial and boundary conditions, the solution can be found.

Let the initial conditions are $d(t_0, x, y) = d_0(x, y)$ and $\dfrac{d(t_0, x, y)}{t} = d_1(x, y)$. Thus, the initial displacement $d_0(x, y)$ and initial velocity $d_1(x, y)$ are given.

Assume that the boundary conditions are

$d(t, x_0, y_0) = 0$ and $d(t, x_f, y_f) = 0$.

Then, the solution is found to be

$$d(t,x,y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} d_{ij}(t,x,y)$$

$$= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \big(A_{ij}\cos l_{ij}t + B_{ij}\sin l_{ij}t\big)\sin\frac{ipx}{a}\sin\frac{jpx}{b},$$

where the eigenvalues (characteristic values) are found as

$$l_{ij} = \sqrt{\frac{T}{r}}p\sqrt{\frac{i^2}{a^2} + \frac{j^2}{b^2}}.$$

Using initial conditions, the Fourier coefficients are obtained in the form of the double Fourier series. In particular, we have

$$A_{ij} = \frac{4}{ab} \int_0^b \int_0^a d_0(x,y) \sin\frac{ipx}{a} \sin\frac{jpy}{b} dxdy ,$$

and $$B_{ij} = \frac{4}{ab\, l_{ij}} \int_0^b \int_0^a d_1(x,y) \sin\frac{ipx}{a} \sin\frac{jpy}{b} dxdy .$$

Example 2.3.6.

Derive the mathematical model of the infinitely long beam on the elastic foundation as shown in Figure 2.3.3. The load force is the square function. The modulus (the spring stiffness per unit length) of the elastic foundation is k_s.

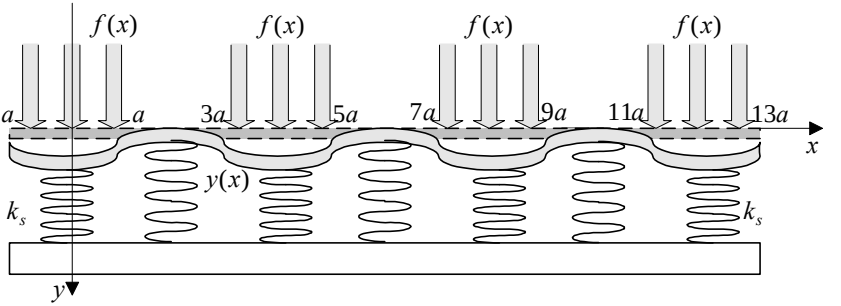


Figure 2.3.3. Beam on elastic foundation under the load force $f(x)$

Solution.

Using the Euler beam theory, the deflection $y(x)$ due to the net load force $F(x)$ is modeled by the fourth-order differential equation

$$k_r \frac{d^4 y}{dt^4} = F(x) ,$$

where k_r is the flexural rigidity constant.

Therefore, we have the following differential equation to model the infinite beam under the consideration

$$k_r \frac{d^4 y}{dt^4} + k_s y = f(x) .$$

The general homogeneous solution is given by

$$y(x) = e^{\frac{1}{2}\sqrt[4]{k_r}xx}\left[k_1 \sin\left(\tfrac{1}{2}\sqrt[4]{k_r}x\right) + k_2 \sin\left(\tfrac{1}{2}\sqrt[4]{k_r}x\right)\right]$$
$$+ e^{\frac{1}{2}\sqrt[4]{k_r}x}\left[k_3 \sin\left(\tfrac{1}{2}\sqrt[4]{k_r}x\right) + k_4 \sin\left(\tfrac{1}{2}\sqrt[4]{k_r}x\right)\right],$$

where the unknown coefficients k_I can be determined using the initial and boundary conditions. The boundary-value problem can be relaxed, and the solution can be found in the series form.

The load force is the periodic function, and using the Fourier series we have

$$f(x) = \frac{f_0}{2} + \frac{2f_0}{p} \sum_{i=1} \frac{\sin\left(\frac{1}{2}ip\right)}{i} \cos\frac{ipx}{2a} .$$

The solution of the differential equation $k_r \dfrac{d^4y}{dt^4} + k_s y = f(x)$ can be found in the following form

$$y(x) = a_0 + \sum_{i=1} a_i \cos\frac{ipx}{2a} .$$

Differentiating this equation four times gives

$$k_s a_0 = \frac{f_0}{2}$$

and $k_r \dfrac{i^4 p^4}{16a^4} + k_s \div a_i = \dfrac{2f_0}{p} \dfrac{\sin(\frac{1}{2}ip)}{i}$.

Thus, the Fourier series coefficients are found as

$$a_0 = \frac{f_0}{2k_s} \text{ and } a_i = \frac{2f_0}{p} \frac{\sin\left(\frac{1}{2}ip\right)}{i\ k_r \left(\frac{ip}{2a}\right)^4 \div + k_s}, \quad i \quad 1.$$

Therefore, the solution is given by

$$y(x) = \frac{f_0}{2k_s} + \frac{32f_0a^4}{p} \sum_{i=1} \frac{\sin\left(\frac{1}{2}ip\right)}{i\left(k_r i^4 p^4 + 16a^4 k_s\right)} \cos\frac{ipx}{2a} .$$

The first-order approximation is

$$y(x) \quad \frac{f_0}{2k} + \frac{32f_0a^4}{p\left(k_r p^4 + 16a^4 k_s\right)} \cos\frac{px}{2a} .$$

Friction Models in Electromechanical Systems

A thorough consideration of friction is essential for understanding the operation of electromechanical systems. Friction is a very complex nonlinear

phenomenon that is difficult to model. The classical *Coulomb* friction is a retarding frictional force (for translational motion) or torque (for rotational motion) that changes its sign with the reversal of the direction of motion, and the amplitude of the frictional force or torque are constant. For translational and rotational motions, the *Coulomb* friction force and torque are

$$F_{Coulomb} = k_{Fc}\,\mathrm{sgn}(v) = k_{Fc}\,\mathrm{sgn}\left(\frac{dx}{dt}\div\right),$$

$$T_{Coulomb} = k_{Tc}\,\mathrm{sgn}(w) = k_{Tc}\,\mathrm{sgn}\left(\frac{dq}{dt}\div\right),$$

where k_{Fc} and k_{Tc} are the Coulomb friction coefficients.

Figure 2.3.4.a illustrates the *Coulomb* friction.

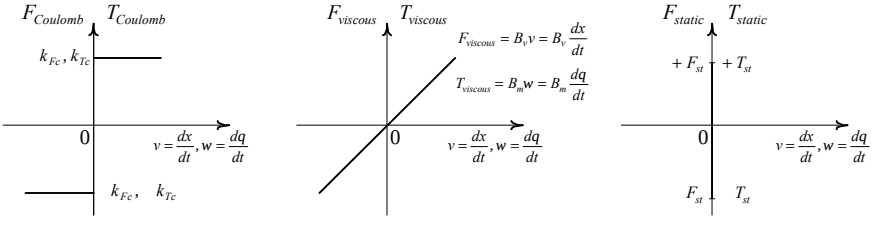


Figure 2.3.4. Functional representations of:
a) *Coulomb* friction; b) viscous friction; c) static friction

Viscous friction is a retarding force or torque that is a linear function of linear or angular velocity. The viscous friction force and torque versus linear and angular velocities are shown in Figure 2.3.4.b. The following expressions are commonly used to model the viscous friction

$$F_{viscous} = B_v v = B_v \frac{dx}{dt} \quad \text{for translational motion,}$$

and $T_{viscous} = B_m w = B_m \frac{dq}{dt}$ for rotational motion,

where B_v and B_m are the viscous friction coefficients.

The static friction exists only when the body is stationary, and vanishes as motion begins. The static friction is a force F_{static} or torque T_{static}, and we have the following expressions

$$F_{static} = \pm F_{st}\Big|_{v=\frac{dx}{dt}=0},$$

and $T_{static} = \pm T_{st}\Big|_{w=\frac{dq}{dt}=0}$.

One concludes that the static friction is a retarding force or torque that tends to prevent the initial translational or rotational motion at the beginning (see Figure 2.3.4.c).

In general, the friction force and torque are nonlinear functions that must be modeled using frictional memory, presliding conditions, etc. The empirical formulas, commonly used to express F_{static} and T_{static}, are

$$F_{fr} = \left(k_{fr1} \quad k_{fr2} e^{\,k|v|} + k_{fr3}|v|\right)\operatorname{sgn}(v_t) = \quad k_{fr1} \quad k_{fr2} e^{\,k\left|\frac{dx}{dt}\right|} + k_{fr3}\left|\frac{dx}{dt}\right| \div \operatorname{sgn} \quad \frac{dx}{dt} \div$$

and

$$T_{fr} = \left(k_{fr1} \quad k_{fr2} e^{\,k|w|} + k_{fr3}|w|\right)\operatorname{sgn}(w) = \quad k_{fr1} \quad k_{fr2} e^{\,k\left|\frac{dq}{dt}\right|} + k_{fr3}\left|\frac{dq}{dt}\right| \div \operatorname{sgn} \quad \frac{dq}{dt} \div$$

These F_{static} and T_{static} are shown in Figure 2.3.5.

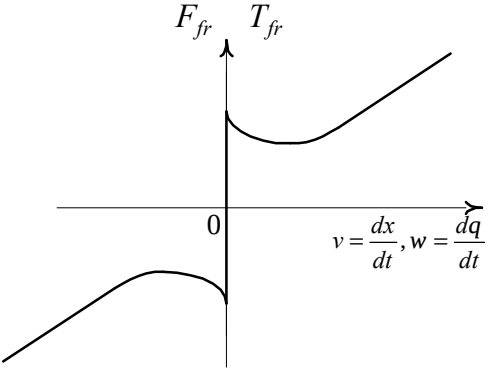


Figure 2.3.5. Friction force and torque are functions of linear and angular velocities

Example 2.3.7. Transducer model

Figure 2.3.6 shows a simple electromechanical device (actuator) with a stationary member and movable plunger. Using Newton's second law, find the differential equations.

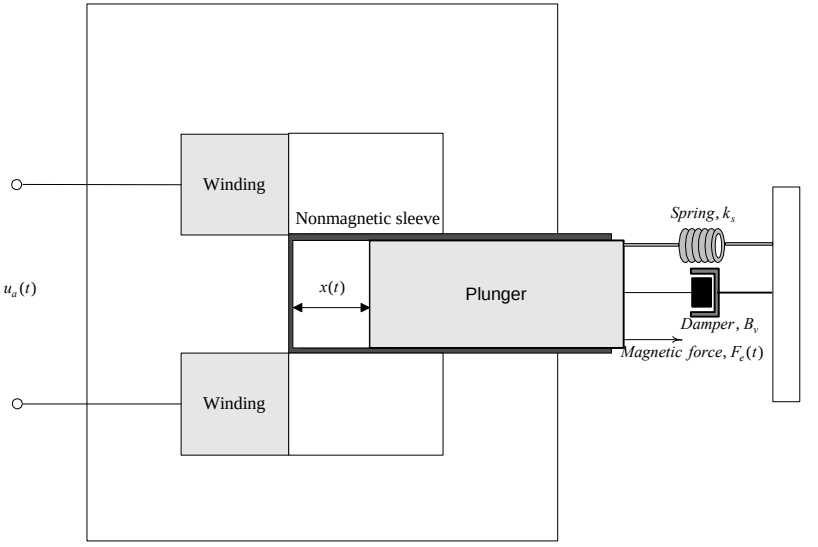


Figure 2.3.6. Schematic of a transducer

Solution.

Let us apply Newton's second law of motion to find the equations of motion and study the dynamics. Newton's law states that the acceleration of an object is proportional to the net force. The vector sum of all forces acting on the object can be found by using a free-body diagram. In particular, for the studied translational mechanical system, one obtains

$$F(t) = m\frac{d^2x}{dt^2} + B_v\frac{dx}{dt} + (k_{s1}x + k_{s2}x^2) + F_e(t),$$

where x denotes the displacement of a plunger; m is the mass of a movable member; B_v is the viscous friction coefficient; k_{s1} and k_{s2} are the spring constants; $F_e(t)$ is the magnetic force,

$$F_e(i,x) = \frac{\P W_c(i,x)}{\P x}.$$

It should be emphasized that Hooke's law is valid only for sufficiently small displacements. The stretch and restoring forces are not directly proportional to the displacement, and these forces are different on either side of the equilibrium position. The restoring/stretching force exerted by the spring is expressed by $(k_{s1}x + k_{s2}x^2)$.

Assuming that the magnetic system is linear, the coenergy is expressed as

$$W_c(i,x) = \tfrac{1}{2}L(x)i^2,$$

then $F_e(i,x) = \frac{1}{2} i^2 \frac{dL(x)}{dx}$.

The inductance is found by using the following formula

$$L(x) = \frac{N^2}{\hat{A}_f + \hat{A}_g} = \frac{N^2 m_f m_0 A_f A_g}{A_g l_f + 2 A_f m_f (x + 2d)},$$

where $\hat{A}_f$ and $\hat{A}_g$ are the reluctances of the ferromagnetic material and air gap; A_f and A_g are the associated cross section areas; l_f and $(x + 2d)$ are the lengths of the magnetic material and the air gap.

Hence, $\frac{dL}{dx} = \frac{2 N^2 m_f^2 m_0 A_f^2 A_g}{[A_g l_f + 2 A_f m_f (x + 2d)]^2}$.

Using Kirchhoff's law, the voltage equation for the electric circuit is given as

$$u_a = ri + \frac{dy}{dt},$$

where the flux linkage y is expressed as $y = L(x)i$.

One obtains

$$u_a = ri + L(x)\frac{di}{dt} + i\frac{dL(x)}{dx}\frac{dx}{dt},$$

and thus

$$\frac{di}{dt} = -\frac{r}{L(x)} i + \frac{2 N^2 m_f^2 m_0 A_f^2 A_g}{L(x)[A_g l_f + 2 A_f m_f (x + 2d)]^2} iv + \frac{1}{L(x)} u_a.$$

Augmenting this equation with differential equation for the mechanical systems

$$F(t) = m\frac{d^2 x}{dt^2} + B_v \frac{dx}{dt} + (k_{s1} x + k_{s2} x^2) + F_e(t),$$

three nonlinear differential equations for the considered transducer are found as

$$\frac{di}{dt} = -\frac{r[A_g l_f + 2 A_f m_f (x + 2d)]}{N^2 m_f m_0 A_f A_g} i + \frac{2 m_f A_f}{A_g l_f + 2 A_f m_f (x + 2d)} iv + \frac{A_g l_f + 2 A_f m_f (x + 2d)}{N^2 m_f m_0 A_f A_g} u_a,$$

$$\frac{dx}{dt} = v,$$

$$\frac{dv}{dt} = \frac{N^2 m_f^2 m_0 A_f^2 A_g}{m[A_g l_f + 2 A_f m_f (x + 2d)]} i^2 - \frac{1}{m}(k_{s1} x + k_{s2} x^2) - \frac{B_v}{m} v.$$

For one-dimensional rotational systems, Newton's second law of motion is expressed as

$$M = Ja \,, \tag{2.3.2}$$

where M is the sum of all moments about the center of mass of a body, (N-m); J is the moment of inertia about its center of mass, (kg-m^2); a is the angular acceleration of the body, (rad/sec^2).

Example 2.3.8.

Given a point mass m suspended by a massless, unstretchable string of length l, (see Figure 2.3.7). Derive the equations of motion for a simple pendulum with negligible friction.

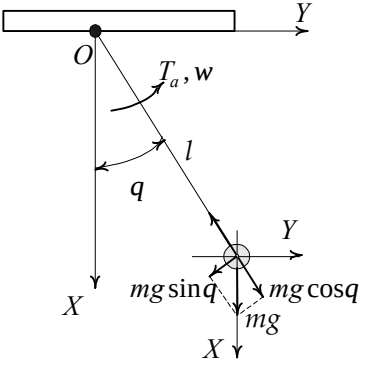


Figure 2.3.7. A simple pendulum

Solution.

The restoring force, which is proportional to $\sin q$ and given by $mg\sin q$, is the tangential component of the net force. Therefore, the sum of the moments about the pivot point O is found as

$$M = \; mgl\sin q + T_a \,,$$

where T_a is the applied torque; l is the length of the pendulum measured from the point of rotation.

Using (2.3.2), one obtains the equation of motion

$$Ja = J\frac{d^2q}{dt^2} = \; mgl\sin q + T_a \,,$$

where J is the moment of inertial of the mass about the point O.

Hence, the second-order differential equation is found to be

$$\frac{d^2q}{dt^2} = \frac{1}{J}\left(\; mgl\sin q + T_a \right).$$

Using the following differential equation for the angular displacement

$$\frac{dq}{dt} = w,$$

one obtains the following set of two first-order differential equations

$$\frac{dw}{dt} = \frac{1}{J}\left(\quad mgl\sin q + T_a \right),$$

$$\frac{dq}{dt} = w.$$

The moment of inertia is expressed by $J = ml^2$. Hence, we have the following differential equations to be used in modeling of a simple pendulum

$$\frac{dw}{dt} = \quad \frac{g}{l}\sin q + \frac{1}{ml^2} T_a,$$

$$\frac{dq}{dt} = w.$$

2.3.2. Lagrange Equations of Motion

Electromechanical systems augment mechanical and electronic components. Therefore, one studies mechanical, electromagnetic, and circuitry transients. It was illustrated that the designer can integrate the *torsional-mechanical* dynamics and circuitry equations of motion. However, there exist general concepts to model systems. The Lagrange and Hamilton concepts are based on the energy analysis. Using the system variables, one finds the total kinetic, dissipation, and potential energies (which are denoted as , D and). Taking note of the total kinetic $t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div,$ dissipation $D\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div,$ and potential $\left(t, q_1,..., q_n\right)$ energies, the Lagrange equations of motion are

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_i}\div\right) - \frac{\P}{\P q_i} + \frac{\P D}{\P\dot{q}_i} + \frac{\P}{\P q_i} = Q_i. \qquad (2.3.3)$$

Here, q_i and Q_i are the generalized coordinates and the generalized forces (applied forces and disturbances). The generalized coordinates q_i are used to derive expressions for energies $t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div,$ $D\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div$ and $\left(t, q_1,..., q_n\right).$

Taking into account that for conservative (losseless) systems $D = 0$, we have the following Lagrange's equations of motion

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_i}\right) - \frac{\partial}{\partial q_i} + \frac{\partial}{\partial q_i} = Q_i .$$

Example 2.3.9. Mathematical model of a simple pendulum

Derive the mathematical model for a simple pendulum using the Lagrange equations of motion.

Solution.

Derivation of the mathematical model for the simple pendulum, shown in Figure 2.3.7, was performed in Example 2.3.8 using the Newtonian mechanics. For the studied conservative (losseless) system we have $D = 0$. Thus, the Lagrange equations of motion are

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_i}\right) - \frac{\partial}{\partial q_i} + \frac{\partial}{\partial q_i} = Q_i .$$

The kinetic energy of the pendulum bob is $= \frac{1}{2} m\left(l\dot{q}\right)^2$.

The potential energy is found as $= mgl\left(1 \quad \cos q\right)$.

As the generalized coordinate, the angular displacement is used, $q_i = q$.

The generalized force is the torque applied, $Q_i = T_a$.

One obtains

$$\frac{\partial}{\partial \dot{q}_i} = \frac{\partial}{\partial \dot{q}} = ml^2\dot{q}\,, \quad \frac{\partial}{\partial q_i} = \frac{\partial}{\partial q} = 0\,, \quad \frac{\partial}{\partial q_i} = \frac{\partial}{\partial q} = mgl\sin q\,.$$

Thus, the first term of the Lagrange equation is found to be

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}}\right) = ml^2\frac{d^2q}{dt^2} + 2ml\frac{dl}{dt}\frac{dq}{dt}.$$

Assuming that the string is unstretchable, we have $\frac{dl}{dt} = 0$.

Hence,

$$ml^2\frac{d^2q}{dt^2} + mgl\sin q = T_a .$$

Thus, one obtains

$$\frac{d^2q}{dt^2} = \frac{1}{ml^2}\left(\quad mgl\sin q + T_a\right).$$

Recall that the equation of motion, derived by using Newtonian mechanics, is

$$\frac{d^2q}{dt^2} = \frac{1}{J}\left(\quad mgl\sin q + T_a \right), \text{ where } J = ml^2 .$$

One concludes that the results are the same, and the equations are

$$\frac{dw}{dt} = \quad \frac{g}{l}\sin q + \frac{1}{ml^2} T_a ,$$

$$\frac{dq}{dt} = w .$$

Example 2.3.10. Mathematical Model of a Pendulum

Consider a double pendulum of two degrees of freedom with no external forces applied to the system (see Figure 2.3.8). Using the Lagrange equations of motion, derive the differential equations.

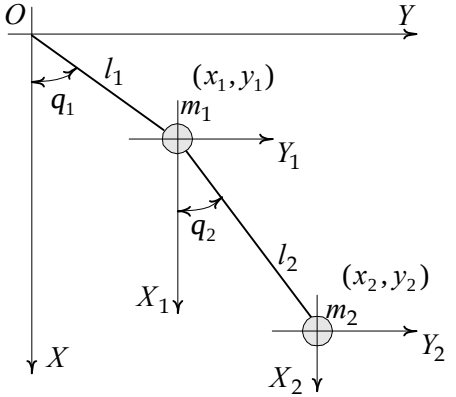


Figure 2.3.8. Double pendulum

Solution.

The angular displacement q_1 and q_2 are chosen as the independent generalized coordinates. In the XY plane studied, let (x_1, y_1) and (x_2, y_2) be the rectangular coordinates of m_1 and m_2. Then, we obtain

$$x_1 = l_1 \cos q_1 , \; x_2 = l_1 \cos q_1 + l_2 \cos q_2 ,$$

$$y_1 = l_1 \sin q_1 , \; y_2 = l_1 \sin q_1 + l_2 \sin q_2 .$$

The total kinetic energy is found to be

$$= \frac{1}{2} m_1 \left(\dot{x}_1^{\,2} + \dot{y}_1^{\,2} \right) + \frac{1}{2} m_2 \left(\dot{x}_2^{\,2} + \dot{y}_2^{\,2} \right)$$

$$= \frac{1}{2}(m_1 + m_2) l_1^2 \dot{q}_1^{\,2} + m_2 l_1 l_2 \dot{q}_1 \dot{q}_2 \cos(q_2 \quad q_1) + \frac{1}{2} m_2 l_2^2 \dot{q}_2^{\,2} .$$

Then, one obtains

$$\frac{\partial}{\partial q_1} = m_2 l_1 l_2 \sin(q_2 - q_1)\dot{q}_1\dot{q}_2,$$

$$\frac{\partial}{\partial \dot{q}_1} = (m_1 + m_2) l_1^2 \dot{q}_1 + m_2 l_1 l_2 \cos(q_2 - q_1)\dot{q}_2,$$

$$\frac{\partial}{\partial q_2} = -m_2 l_1 l_2 \sin(q_1 - q_2)\dot{q}_1\dot{q}_2,$$

$$\frac{\partial}{\partial \dot{q}_2} = m_2 l_1 l_2 \cos(q_2 - q_1)\dot{q}_1 + m_2 l_1^2 \dot{q}_2.$$

The total potential energy is given by

$$= m_1 g y_1 + m_2 g y_2 = -(m_1 + m_2) g l_1 \cos q_1 - m_2 g l_2 \cos q_2.$$

Hence, $\frac{\partial}{\partial q_1} = (m_1 + m_2) g l_1 \sin q_1$ and $\frac{\partial}{\partial q_2} = m_2 g l_2 \sin q_2$.

The Lagrange equations of motion are

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_1}\right) - \frac{\partial}{\partial q_1} + \frac{\partial}{\partial q_1} = 0,$$

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_2}\right) - \frac{\partial}{\partial q_2} + \frac{\partial}{\partial q_2} = 0.$$

Hence, the dynamic equations of the system are

$$(m_1 + m_2) l_1 \ddot{q}_1 + m_2 l_2 \cos(q_2 - q_1)\ddot{q}_2 - m_2 l_2 \sin(q_2 - q_1)\dot{q}_2^2 + (m_1 + m_2) g \sin q_1 = 0,$$

$$l_2 \ddot{q}_2 + l_1 \cos(q_2 - q_1)\ddot{q}_1 + l_1 \sin(q_2 - q_1)\dot{q}_1^2 + g \sin q_2 = 0.$$

It should be emphasized that if the torques T_1 and T_2 are applied to the first and second joints, the following equations of motions results

$$(m_1 + m_2) l_1 \ddot{q}_1 + m_2 l_2 \cos(q_2 - q_1)\ddot{q}_2 - m_2 l_2 \sin(q_2 - q_1)\dot{q}_2^2 + (m_1 + m_2) g \sin q_1 = T_1,$$

$$l_2 \ddot{q}_2 + l_1 \cos(q_2 - q_1)\ddot{q}_1 + l_1 \sin(q_2 - q_1)\dot{q}_1^2 + g \sin q_2 = T_2.$$

Example 2.3.11. Mathematical Model of a Circuit Network

Consider a two-mesh electric circuit, as shown in Figure 2.3.9. Find the circuitry dynamics.

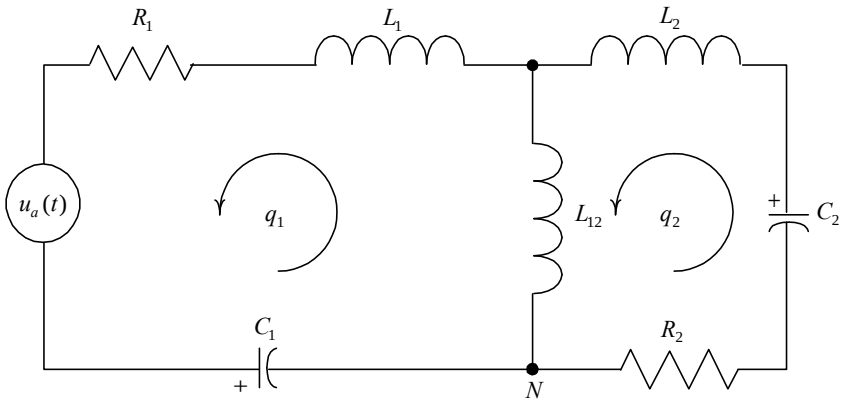


Figure 2.3.9. Two-mesh circuit network

Solution.

We use q_1 and q_2 as the independent generalized coordinates, where q_1 is the electric charge in the first loop, q_2 represents the electric charge in the second loop. The generalized force, which is applied to the system, is denoted as Q_1. These generalized coordinates are related to the circuitry variables. In particular, the currents i_1 and i_2 are found in terms of charges, $i_1 = \dot{q}_1$ and $i_2 = \dot{q}_2$. That is, $q_1 = \frac{i_1}{s}$ and $q_2 = \frac{i_2}{s}$. The generalized force is the applied voltage, $u_a(t) = Q_1$.

The total magnetic energy (kinetic energy) is expressed by

$$= \tfrac{1}{2} L_1 \dot{q}_1^{\,2} + \tfrac{1}{2} L_{12}\left(\dot{q}_1 \quad \dot{q}_2\right)^2 + \tfrac{1}{2} L_2 \dot{q}_2^{\,2}.$$

By using this equation for , we have

$$\frac{\P}{\P q_1} = 0, \ \frac{\P}{\P \dot{q}_1} = \left(L_1 + L_{12}\right)\dot{q}_1 \quad L_{12}\dot{q}_2,$$

$$\frac{\P}{\P q_2} = 0, \ \frac{\P}{\P \dot{q}_2} = \quad L_{12}\dot{q}_1 + \left(L_2 + L_{12}\right)\dot{q}_2.$$

Using the equation for the total electric energy (potential energy)

$$= \tfrac{1}{2}\frac{q_1^2}{C_1} + \tfrac{1}{2}\frac{q_2^2}{C_2},$$

one finds

$$\frac{\P}{\P q_1} = \frac{q_1}{C_1} \text{ and } \frac{\P}{\P q_2} = \frac{q_2}{C_2}.$$

The total heat energy dissipated is

$$D = \tfrac{1}{2} R_1 \dot{q}_1^{\,2} + \tfrac{1}{2} R_2 \dot{q}_2^{\,2}.$$

Hence,

$$\frac{\partial D}{\partial \dot{q}_1} = R_1 \dot{q}_1 \text{ and } \frac{\partial D}{\partial \dot{q}_2} = R_2 \dot{q}_2.$$

The Lagrange equations of motion are expressed using the independent coordinates used. We obtain

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_1}\right) - \frac{\partial}{\partial q_1} + \frac{\partial D}{\partial \dot{q}_1} + \frac{\partial}{\partial q_1} = Q_1,$$

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_2}\right) - \frac{\partial}{\partial q_2} + \frac{\partial D}{\partial \dot{q}_2} + \frac{\partial}{\partial q_2} = 0.$$

Hence, the differential equations for the circuit studied are found to be

$$\left(L_1 + L_{12}\right)\ddot{q}_1 \quad L_{12}\ddot{q}_2 + R_1\dot{q}_1 + \frac{q_1}{C_1} = u_a,$$

$$L_{12}\ddot{q}_1 + \left(L_2 + L_{12}\right)\ddot{q}_2 + R_2\dot{q}_2 + \frac{q_2}{C_2} = 0.$$

The SIMULINK model can be built using these derived nonlinear differential equations. In particular, we have

$$\ddot{q}_1 = \frac{1}{\left(L_1 + L_{12}\right)} \quad \frac{q_1}{C_1} \quad R_1\dot{q}_1 + L_{12}\ddot{q}_2 + u_a \div$$

and $\ddot{q}_2 = \dfrac{1}{\left(L_2 + L_{12}\right)} \quad L_{12}\ddot{q}_1 \quad \dfrac{q_2}{C_2} \quad R_2\dot{q}_2 \div .$

The corresponding SIMULINK diagram is shown in Figure 2.3.10.

It should be emphasized that the currents i_1 and i_2 are expressed in terms of charges as

$i_1 = \dot{q}_1$ and $i_2 = \dot{q}_2$.

That is, we have

$$q_1 = \frac{i_1}{s} \text{ and } q_2 = \frac{i_2}{s}.$$

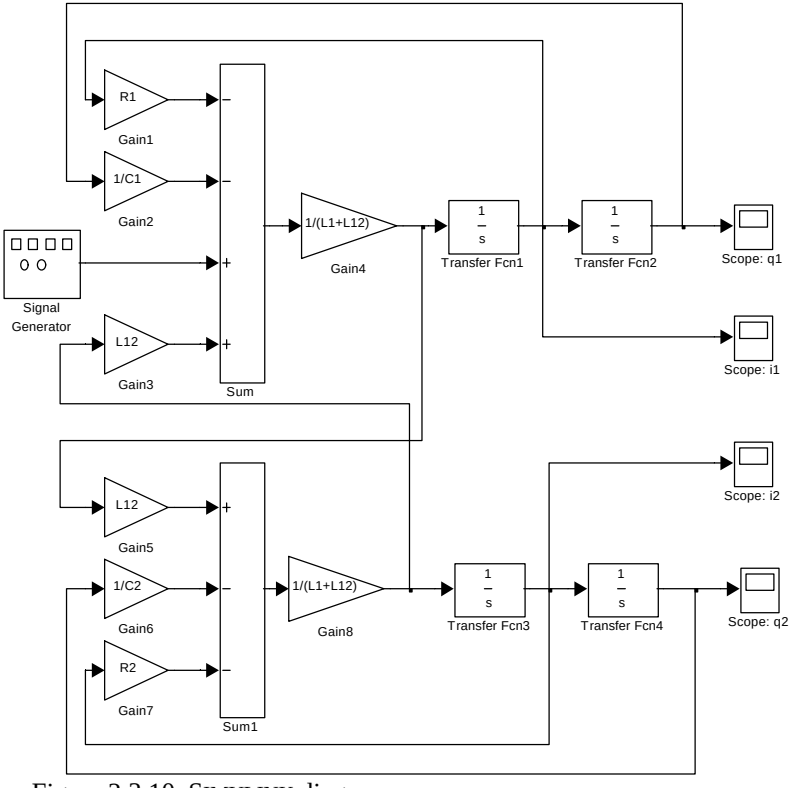
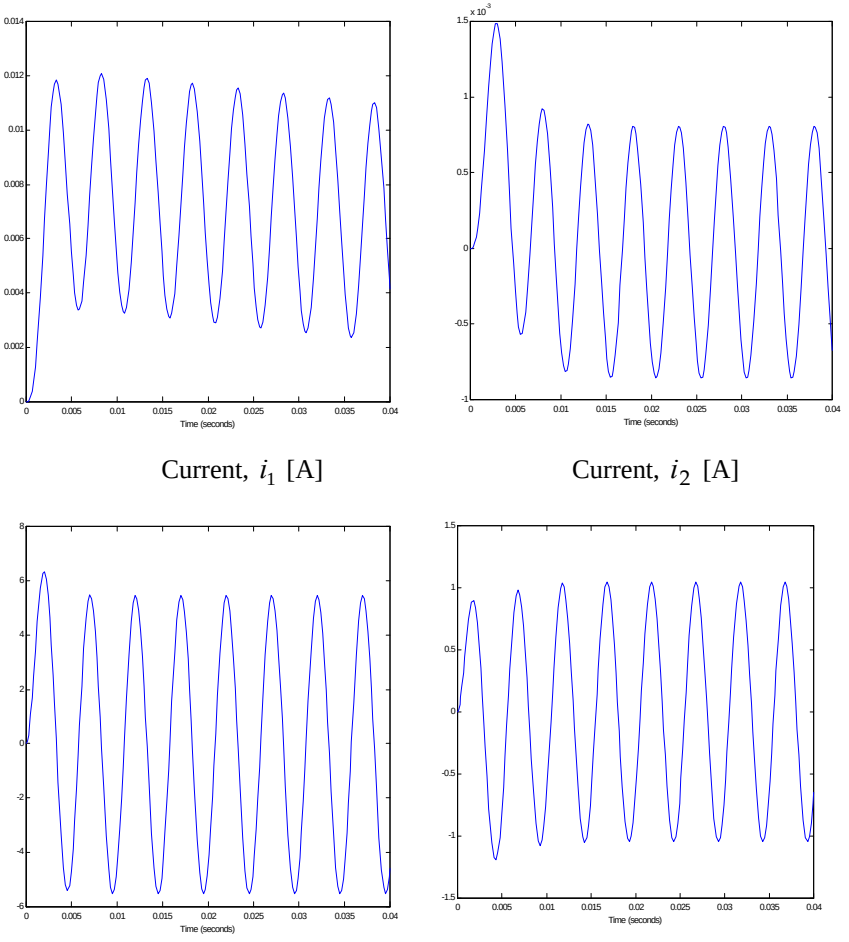


Figure 2.3.10. SIMULINK diagram

To perform simulations (numerical analysis), one must use the parameter values. The circuitry parameters are assigned to be: L_1 =0.01 H, L_2 =0.005 H, L_{12} = 0.0025 H, C_1 =0.02 F, C_2 =0.1 F, R_1 =10 ohm, R_2 = 5 ohm and $u_a = 100\sin(200t)$ V.

Simulation results, which give the time history of $q_1(t), q_2(t), i_1(t)$ and $i_2(t)$, are documented in Figure 2.3.11.

Generalized coordinate, q_1 Generalized coordinate, q_2



Current, i_1 [A] Current, i_2 [A]

Figure 2.3.11. Circuit dynamics: evolution of the generalized coordinates and currents

Example 2.3.12. Mathematical Model of an Electric Circuit

Using the Lagrange equations of motion, develop the mathematical models for the circuit shown in Figure 2.3.12. Prove that the model derived using the Lagrange equations of motion are equivalent to the model developed using Kirchhoff's law.

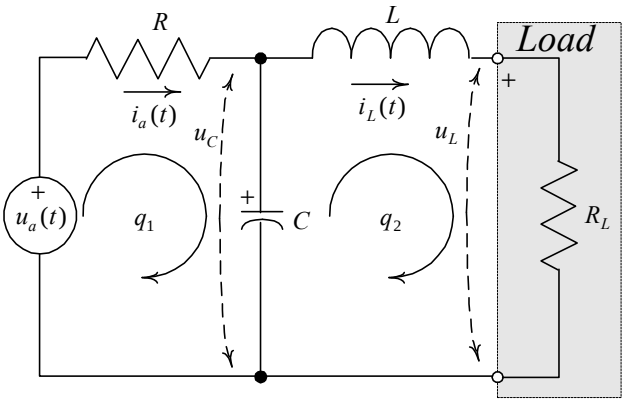


Figure 2.3.12. Electric circuit

Solution.

Using q_1 and q_2 as the independent generalized coordinates, the Lagrange equations of motion can be found. Here, q_1 is the electric charge in the first loop and $i_a = \dot{q}_1$, and q_2 is the electric charge in the second loop, $i_L = \dot{q}_2$. The generalized force, applied to the system, is denoted as Q_1, and $u_a(t) = Q_1$.

The total kinetic energy is $= \frac{1}{2} L \dot{q}_2^{\,2}$.

Therefore, we have,

$$\frac{\P}{\P q_1} = 0, \ \frac{\P}{\P \dot{q}_1} = 0, \text{ and } \frac{d}{dt}\left(\frac{\P}{\P \dot{q}_1}\right) = 0,$$

$$\frac{\P}{\P q_2} = 0, \ \frac{\P}{\P \dot{q}_2} = L\dot{q}_2, \text{ and } \frac{d}{dt}\left(\frac{\P}{\P \dot{q}_2}\right) = L\ddot{q}_2.$$

The total potential energy is expressed as

$$= \frac{1}{2}\frac{(q_1 \quad q_2)^2}{C}.$$

Hence

$$\frac{\P}{\P q_1} = \frac{q_1 \quad q_2}{C} \text{ and } \frac{\P}{\P q_2} = \frac{q_1 + q_2}{C}.$$

The total dissipated energy is

$$D = \frac{1}{2} R \dot{q}_1^{\,2} + \frac{1}{2} R_L \dot{q}_2^{\,2}.$$

Therefore

$$\frac{\P D}{\P \dot{q}_1} = R\dot{q}_1 \text{ and } \frac{\P D}{\P \dot{q}_2} = R_L \dot{q}_2 .$$

The Lagrange equations of motion

$$\frac{d}{dt}\left(\frac{\P}{\P \dot{q}_1}\right) - \frac{\P}{\P q_1} + \frac{\P D}{\P \dot{q}_1} + \frac{\P}{\P q_1} = Q_1 ,$$

$$\frac{d}{dt}\left(\frac{\P}{\P \dot{q}_2}\right) - \frac{\P}{\P q_2} + \frac{\P D}{\P \dot{q}_2} + \frac{\P}{\P q_2} = 0$$

lead one to the following two differential equations

$$R\dot{q}_1 + \frac{q_1 - q_2}{C} = u_a ,$$

$$L\ddot{q}_2 + R_L \dot{q}_2 + \frac{q_1 + q_2}{C} = 0 .$$

Hence, we have found a set of two differential equations. In particular,

$$\dot{q}_1 = \frac{1}{R}\left(-\frac{q_1 + q_2}{C} + u_a\right),$$

$$\ddot{q}_2 = \frac{1}{L}\left(-R_L \dot{q}_2 + \frac{q_1 - q_2}{C}\right).$$

By using Kirchhoff's law, two differential equations result

$$\frac{du_C}{dt} = \frac{1}{C}\left(\frac{u_C}{R} - i_L + \frac{u_a(t)}{R}\right),$$

$$\frac{di_L}{dt} = \frac{1}{L}\left(u_C - R_L i_L\right).$$

Taking note of $i_a = \dot{q}_1$ and $i_L = \dot{q}_2$, and making use $C\frac{du_C}{dt} = i_a - i_L$, we obtain

$$u_C = \frac{q_1 - q_2}{C} .$$

The equivalence of the differential equations derived using the Lagrange equations of motion and Kirchhoff's law is proven.

Example 2.3.13. Mathematical model of a boost converter

A high-frequency, one-quadrant *boost* (*step-up*) dc-dc switching converter is documented in Figure 2.3.13. Find the mathematical model in the form of differential equations.

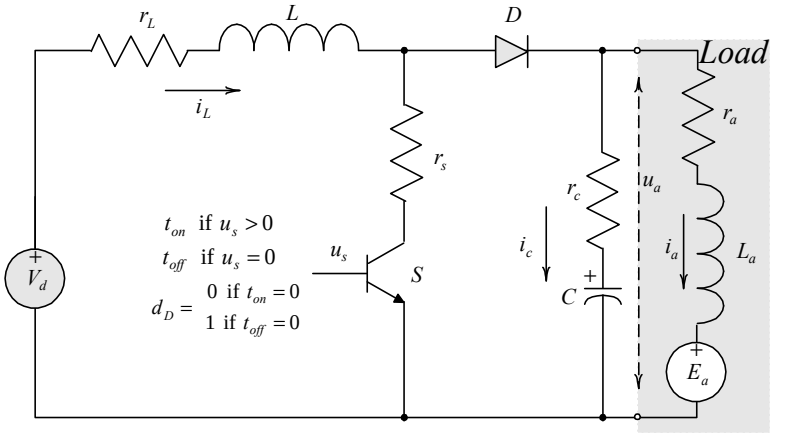


Figure 2.3.13. *Boost* converter

Solution.

To solve the model development problem, we will derive the differential equations if the duty ratio d_D is 1 and 0. Then, we will augment two mathematical models found to model the *boost* converter.

When the switch is closed, the diode *D* is reverse biased. For $d_D = 1$ (t_{off} = 0), one obtains the following set of linear differential equations

$$\frac{du_C}{dt} = \frac{1}{C} i_a,$$

$$\frac{di_L}{dt} = \frac{1}{L}\left(-(r_L + r_s) i_L + V_d \right),$$

$$\frac{di_a}{dt} = \frac{1}{L_a}\left(u_C - (r_a + r_c) i_a - E_a \right).$$

If the switch is open ($d_D = 0$), the diode *D* is forward biased because the direction of the inductor current i_L does not change instantly. Therefore, one has three linear differential equations

$$\frac{du_C}{dt} = \frac{1}{C}\left(i_L - i_a \right),$$

$$\frac{di_L}{dt} = \frac{1}{L}\left(-u_C - (r_L + r_c) i_L + r_c i_a + V_d \right),$$

$$\frac{di_a}{dt} = \frac{1}{L_a}\left(u_C + r_c i_L - (r_a + r_c) i_a - E_a \right).$$

Assuming the switching frequency id high, the *averaging* concept is applied, and we have

$$\frac{du_C}{dt} = \frac{1}{C}\left(i_L \quad i_a \quad i_L d_D\right),$$

$$\frac{di_L}{dt} = \frac{1}{L}\left(\quad u_C \quad \left(r_L + r_c\right)i_L + r_c i_a + u_C d_D + \left(r_c \quad r_s\right)i_L d_D \quad r_c i_a d_D + V_d\right),$$

$$\frac{di_a}{dt} = \frac{1}{L_a}\left(u_C + r_c i_L \quad \left(r_a + r_c\right)i_a \quad r_c i_L d_D \quad E_a\right).$$

Considering the duty ratio as the control input, one concludes that a set of nonlinear differential equations result. In fact, the state variables are multiplied by the control.

Let us illustrate that Lagrange's concept gives the same differential equations. We denote the electric charges in the first and the second loops as q_1 and q_2, and the generalized forces are Q_1 and Q_2. Then,

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_1}\right) - \frac{\P}{\P q_1} + \frac{\P D}{\P\dot{q}_1} + \frac{\P}{\P q_1} = Q_1 \ ,$$

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_2}\right) - \frac{\P}{\P q_2} + \frac{\P D}{\P\dot{q}_2} + \frac{\P}{\P q_2} = Q_2 .$$

For the closed switch, the total kinetic, potential, and dissipated energies are

$$= \tfrac{1}{2}\left(L\dot{q}_1^{\,2} + L_a\dot{q}_2^{\,2}\right), \quad = \tfrac{1}{2}\frac{q_2^2}{C}, D = \tfrac{1}{2}\left(\left(r_L + r_s\right)\dot{q}_1^{\,2} + \left(r_c + r_a\right)\dot{q}_2^{\,2}\right).$$

Assuming that the resistances, inductances, and capacitance are time-invariant (constant), one obtains

$$\frac{\P}{\P q_1} = 0 \ , \ \frac{\P}{\P q_2} = 0 \ , \ \frac{\P}{\P\dot{q}_1} = L\dot{q}_1 \ , \ \frac{\P}{\P\dot{q}_2} = L_a\dot{q}_2 \ ,$$

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_1}\right) = L\ddot{q}_1 \ , \frac{d}{dt}\left(\frac{\P}{\P\dot{q}_2}\right) = L_a\ddot{q}_2 \ ,$$

$$\frac{\P}{\P q_1} = 0 \ , \ \frac{\P}{\P q_2} = \frac{q_2}{C} \ ,$$

$$\frac{\P D}{\P\dot{q}_1} = \left(r_L + r_s\right)\dot{q}_1 \ , \ \frac{\P D}{\P\dot{q}_2} = \left(r_c + r_a\right)\dot{q}_2 .$$

Therefore,

$$L\ddot{q}_1 + \left(r_L + r_s\right)\dot{q}_1 = Q_1 ,$$

$$L_a\ddot{q}_2 + \left(r_c + r_a\right)\dot{q}_2 + \frac{1}{C}q_2 = Q_2 \ ,$$

and thus,

$$\ddot{q}_1 = \frac{1}{L}\Big(\quad (r_L + r_s)\dot{q}_1 + Q_1\Big),$$

$$\ddot{q}_2 = \frac{1}{L_a} \quad (r_c + r_a)\dot{q}_2 \quad \frac{1}{C}q_2 + Q_2 \div,$$

The total kinetic, potential, and dissipated energies if the switch is open are found to be

$$= \tfrac{1}{2}\Big(L\dot{q}_1{}^2 + L_a\dot{q}_2{}^2\Big), \quad = \tfrac{1}{2}\frac{(q_1 \quad q_2)^2}{C}, D = \tfrac{1}{2}\Big(r_L\dot{q}_1{}^2 + r_c(\dot{q}_1 \quad \dot{q}_2)^2 + r_a\dot{q}_2{}^2\Big).$$

Thus,

$$\frac{\P}{\P q_1} = 0, \ \frac{\P}{\P q_2} = 0, \ \frac{\P}{\P \dot{q}_1} = L\dot{q}_1, \ \frac{\P}{\P \dot{q}_2} = L_a\dot{q}_2,$$

$$\frac{d}{dt}\ \frac{\P}{\P \dot{q}_1} \div = L\ddot{q}_1, \ \frac{d}{dt}\ \frac{\P}{\P \dot{q}_2} \div = L_a\ddot{q}_2,$$

$$\frac{\P}{\P q_1} = \frac{q_1 \quad q_2}{C}, \ \frac{\P}{\P q_2} = \quad \frac{q_1 \quad q_2}{C},$$

$$\frac{\P D}{\P \dot{q}_1} = (r_L + r_c)\dot{q}_1 \quad r_c\dot{q}_2, \ \frac{\P D}{\P \dot{q}_2} = -r_c\dot{q}_1 + (r_c + r_a)\dot{q}_2.$$

Using

$$L\ddot{q}_1 + (r_L + r_c)\dot{q}_1 \quad r_c\dot{q}_2 + \frac{q_1 \quad q_2}{C} = Q_1,$$

$$L_a\ddot{q}_2 \quad r_c\dot{q}_1 + (r_c + r_a)\dot{q}_2 \quad \frac{q_1 \quad q_2}{C} = Q_2,$$

one has

$$\ddot{q}_1 = \frac{1}{L} \quad (r_L + r_c)\dot{q}_1 + r_c\dot{q}_2 \quad \frac{q_1 \quad q_2}{C} + Q_1 \div,$$

$$\ddot{q}_2 = \frac{1}{L_a} \ r_c\dot{q}_1 \quad (r_c + r_a)\dot{q}_2 + \frac{q_1 \quad q_2}{C} + Q_2 \div.$$

It must be emphasized that $i_L = \dot{q}_1$, $i_a = \dot{q}_2$, and $Q_1 = V_d$, $Q_2 = \quad E_a$. Taking note of the differential equations when the switch is closed and open, the differential equations in Cauchy's form are found using $\frac{dq_1}{dt} = i_L$ and $\frac{dq_2}{dt} = i_a$. The voltage across the capacitor u_C is expressed using the charges q_1 and q_2. When the switch is closed $u_C = \ \frac{q_2}{C}$. If the switch is

open $u_C = \dfrac{q_1 \quad q_2}{C}$. The analysis of the differential equations derived using Kirchhoff's voltage law and the Lagrange equations of motion illustrates that the mathematical models are found using different state variables. In particular, u_C, i_L, i_a and q_1, i_L, q_2, i_a are used. However, the resulting differential equations are the same as one applies the corresponding variable transformations as given by

$$\frac{dq_1}{dt} = i_L, \ \frac{dq_2}{dt} = i_a, \ Q_1 = V_d \text{ and } Q_2 = \quad E_a.$$

Example 2.3.14. Mathematical model of an electric motor

Consider a motor with two independently excited stator and rotor windings, see Figure 2.3.14. Derive the differential equations.

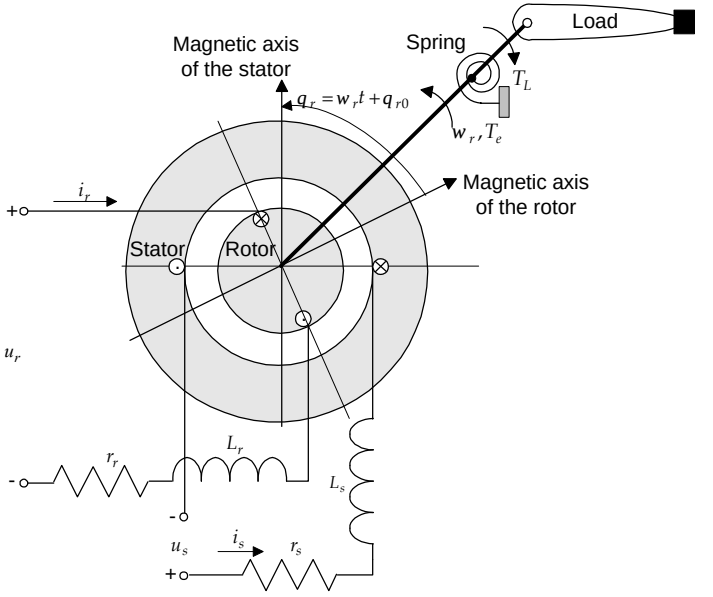


Figure 2.3.14. Motor with stator and rotor windings

Solution.

The following notations are used: i_s and i_r are the currents in the stator and rotor windings; u_s and u_r are the applied voltages to the stator and rotor windings; w_r and q_r are the rotor angular velocity and displacement; T_e and T_L are the electromagnetic and load torques; r_s and r_r are the resistances of the stator and rotor windings; L_s and L_r are the self-inductances of the stator and rotor windings; L_{sr} is the mutual inductance of

the stator and rotor windings; $_m$ is the reluctance of the magnetizing path; N_s and N_r are the number of turns in the stator and rotor windings; J is the moment of inertia of the rotor and attached load; B_m is the viscous friction coefficient; k_s is the spring constant.

The magnetic fluxes that cross an air gap produce a force of attraction, and the developed electromagnetic torque T_e is countered by the tortional spring which causes a *counterclockwise* rotation. The load torque T_L should be considered.

Our goal is to find a nonlinear mathematical model. In fact, the ability to formulate the modeling problem and find the resulting equations that describe a motion device constitute the most important issues. By using the Lagrange concept, the independent generalized coordinates must be chosen. Let us use q_1, q_2 and q_3, where q_1 and q_2 denote the electric charges in the stator and rotor windings; q_3 represents the rotor angular displacement.

We denote the generalized forces, applied to an electromechanical system, as Q_1, Q_2 and Q_3, where Q_1 and Q_2 are the applied voltages to the stator and rotor windings; Q_3 is the load torque.

The first derivative of the generalized coordinates $\dot{q}_1$ and $\dot{q}_2$ represent the stator and rotor currents i_s and i_r, while $\dot{q}_3$ is the angular velocity of the rotor w_r. We have,

$$q_1 = \frac{i_s}{s},\ q_2 = \frac{i_r}{s},\ q_3 = q_r,\ \dot{q}_1 = i_s,\ \dot{q}_2 = i_r,\ \dot{q}_3 = w_r,$$

$$Q_1 = u_s,\ Q_2 = u_r \text{ and } Q_3 = \ T_L.$$

The Lagrange equations are expressed in terms of each independent coordinates, and we have

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_1}\right) - \frac{\P}{\P q_1} + \frac{\P D}{\P\dot{q}_1} + \frac{\P}{\P q_1} = Q_1,$$

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_2}\right) - \frac{\P}{\P q_2} + \frac{\P D}{\P\dot{q}_2} + \frac{\P}{\P q_2} = Q_2,$$

$$\frac{d}{dt}\left(\frac{\P}{\P\dot{q}_3}\right) - \frac{\P}{\P q_3} + \frac{\P D}{\P\dot{q}_3} + \frac{\P}{\P q_3} = Q_3.$$

The total kinetic energy of electrical and mechanical systems is found as a sum of the total magnetic (electrical) $_E$ and mechanical $_M$ energies. The total kinetic energy of the stator and rotor circuitry is given as

$$_E = \tfrac{1}{2} L_s \dot{q}_1^2 + L_{sr}\dot{q}_1\dot{q}_2 + \tfrac{1}{2} L_r \dot{q}_2^2.$$

The total kinetic energy of the mechanical system, which is a function of the equivalent moment of inertia of the rotor and the payload attached, is expressed by

$$_M = \tfrac{1}{2} J\dot{q}_3^2 .$$

Then, we have

$$= \ _E + \ _M = \tfrac{1}{2} L_s\dot{q}_1^2 + L_{sr}\dot{q}_1\dot{q}_2 + \tfrac{1}{2} L_r\dot{q}_2^2 + \tfrac{1}{2} J\dot{q}_3^2 .$$

The mutual inductance is a periodic function of the angular rotor displacement, and $L_{sr}(q_r) = \dfrac{N_s N_r}{_m(q_r)}$.

The magnetizing reluctance is maximum if the stator and rotor windings are not displaced, and $_m(q_r)$ is minimum if the coils are displaced by 90 degrees. Then, $L_{sr\min} \quad L_{sr}(q_r) \quad L_{sr\max}$, where $L_{sr\max} = \dfrac{N_s N_r}{_m(90^\circ)}$ and $L_{sr\min} = \dfrac{N_s N_r}{_m(0^\circ)}$.

The mutual inductance can be approximated as a cosine function of the rotor angular displacement. The amplitude of the mutual inductance between the stator and rotor windings is found as $L_M = L_{sr\max} = \dfrac{N_s N_r}{_m(90^\circ)}$.

Then,

$$L_{sr}(q_r) = L_M \cos q_r = L_M \cos q_3 .$$

One obtains an explicit expression for the total kinetic energy as

$$= \tfrac{1}{2} L_s\dot{q}_1^2 + L_M\dot{q}_1\dot{q}_2 \cos q_3 + \tfrac{1}{2} L_r\dot{q}_2^2 + \tfrac{1}{2} J\dot{q}_3^2 .$$

The following partial derivatives result

$$\frac{\P}{\P q_1} = 0 , \ \frac{\P}{\P\dot{q}_1} = L_s\dot{q}_1 + L_M\dot{q}_2 \cos q_3 ,$$

$$\frac{\P}{\P q_2} = 0 , \ \frac{\P}{\P\dot{q}_2} = L_M\dot{q}_1 \cos q_3 + L_r\dot{q}_2 ,$$

$$\frac{\P}{\P q_3} = \ L_M\dot{q}_1\dot{q}_2 \sin q_3 , \ \frac{\P}{\P\dot{q}_3} = J\dot{q}_3 .$$

The potential energy of the spring with constant k_s is

$$= \tfrac{1}{2} k_s q_3^2 .$$

Therefore,

$$\frac{\P}{\P q_1} = 0 , \ \frac{\P}{\P q_2} = 0 , \text{ and } \frac{\P}{\P q_3} = k_s q_3 .$$

The total heat energy dissipated is expressed as

$$D = D_E + D_M,$$

where D_E is the heat energy dissipated in the stator and rotor windings, $D_E = \frac{1}{2} r_s \dot{q}_1^2 + \frac{1}{2} r_r \dot{q}_2^2$; D_M is the heat energy dissipated by mechanical system, $D_M = \frac{1}{2} B_m \dot{q}_3^2$.

Hence,

$$D = \frac{1}{2} r_s \dot{q}_1^2 + \frac{1}{2} r_r \dot{q}_2^2 + \frac{1}{2} B_m \dot{q}_3^2.$$

One obtains

$$\frac{\P D}{\P \dot{q}_1} = r_s \dot{q}_1, \quad \frac{\P D}{\P \dot{q}_2} = r_r \dot{q}_2 \text{ and } \frac{\P D}{\P \dot{q}_3} = B_m \dot{q}_3.$$

Using

$$q_1 = \frac{i_s}{s}, \; q_2 = \frac{i_r}{s}, \; q_3 = q_r, \; \dot{q}_1 = i_s, \; \dot{q}_2 = i_r, \; \dot{q}_3 = w_r,$$

$$Q_1 = u_s, \; Q_2 = u_r \text{ and } Q_3 = \quad T_L,$$

we have three differential equations for a servo-system. In particular,

$$L_s \frac{di_s}{dt} + L_M \cos q_r \frac{di_r}{dt} \quad L_M i_r \sin q_r \frac{dq_r}{dt} + r_s i_s = u_s,$$

$$L_r \frac{di_r}{dt} + L_M \cos q_r \frac{di_s}{dt} \quad L_M i_s \sin q_r \frac{dq_r}{dt} + r_r i_r = u_r,$$

$$J \frac{d^2 q_r}{dt^2} + L_M i_s i_r \sin q_r + B_m \frac{dq_r}{dt} + k_s q_r = \quad T_L.$$

The last equation should be rewritten by making use the rotor angular velocity; that is,

$$\frac{dq_r}{dt} = w_r.$$

Finally, using the stator and rotor currents, angular velocity and position as the state variables, the nonlinear differential equations in Cauchy's form are found as

$$\frac{di_s}{dt} = \frac{r_s L_r i_s \quad \frac{1}{2} L_M^2 i_s w_r \sin 2q_r + r_r L_M i_r \cos q_r + L_r L_M i_r w_r \sin q_r + L_r u_s \quad L_M \cos q_r u_r}{L_s L_r \quad L_M^2 \cos^2 q_r},$$

$$\frac{di_r}{dt} = \frac{r_s L_M i_s \cos q_r + L_s L_M i_s w_r \sin q_r \quad r_r L_s i_r \quad \frac{1}{2} L_M^2 i_r w_r \sin 2q_r \quad L_M \cos q_r u_s + L_s u_r}{L_s L_r \quad L_M^2 \cos^2 q_r},$$

$$\frac{dw_r}{dt} = \frac{1}{J}\left(\quad L_M i_s i_r \sin q_r \quad B_m w_r \quad k_s q_r \quad T_L \right),$$

$$\frac{dq_r}{dt} = w_r.$$

The developed nonlinear mathematical model in the form of highly coupled nonlinear differential equations cannot be linearized, and one must model the doubly exited transducer studied using the nonlinear differential equations derived.

2.3.3. Hamilton Equations of Motion

The Hamilton concept allows one to model the system dynamics, and the differential equations are found using the generalized momenta p_i, $p_i = \frac{L}{\dot{q}_i}$ (the generalized coordinates were used in the Lagrange equations of motion).

The Lagrangian function $L\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div$ for the conservative systems is the difference between the total kinetic and potential energies. In particular,

$$L\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div = \quad t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div \quad (t, q_1,..., q_n)\ .$$

Thus, $L\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div$ is the function of $2n$ independent variables. One has

$$dL = \sum_{i=1}^{n} \frac{L}{q_i} dq_i + \frac{L}{\dot{q}_i} d\dot{q}_i \div = \sum_{i=1}^{n} \left(\dot{p}_i dq_i + p_i d\dot{q}_i\right).$$

We define the Hamiltonian function as

$$H(t, q_1,..., q_n, p_1,..., p_n) = \quad L\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div + \sum_{i=1}^{n} p_i \dot{q}_i\ ,$$

$$dH = \sum_{i=1}^{n} \left(\dot{p}_i dq_i + \dot{q}_i dp_i \right),$$

where $\sum_{i=1}^{n} p_i \dot{q}_i = \sum_{i=1}^{n} \frac{L}{\dot{q}_i} \dot{q}_i = \sum_{i=1}^{n} \frac{}{\dot{q}_i} \dot{q}_i = 2 \quad .$

Thus, we have

$$H\ t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div = \quad t, q_1,..., q_n, \frac{dq_1}{dt},..., \frac{dq_n}{dt}\div + \quad (t, q_1,..., q_n)$$

or $H(t, q_1,..., q_n, p_1,..., p_n) = (t, q_1,..., q_n, p_1,..., p_n) + \quad (t, q_1,..., q_n).$

One concludes that the Hamiltonian, which is equal to the total energy, is expressed as a function of the generalized coordinates and generalized momenta. The equations of motion are governed by the following equations

$$\dot{p}_i = \frac{H}{q_i}, \ \dot{q}_i = \frac{H}{p_i}, \tag{2.3.4}$$

which are called the Hamiltonian equations of motion.

It is evident that using the Hamiltonian mechanics, one obtains the system of $2n$ first-order partial differential equations to model the system dynamics. In contrast, using the Lagrange equations of motion, the system of n second-order differential equations results. However, the derived differential equations are equivalent.

Example 2.3.15.

Consider the harmonic oscillator. The total energy is given as the sum of the kinetic and potential energies, $_T = \ + \ = \frac{1}{2}(mv^2 + k_s x^2)$. Find the equations of motion using the Lagrange and Hamilton concepts.

Solution.

The Lagrangian function is

$$L \ x, \frac{dx}{dt} \div = \qquad = \frac{1}{2}(mv^2 \quad k_s x^2) = \frac{1}{2}(m\dot{x}^2 \quad k_s x^2).$$

Making use of the Lagrange equations of motion

$$\frac{d}{dt}\frac{L}{\dot{x}} \quad \frac{L}{x} = 0,$$

we have

$$m\frac{d^2x}{dt^2} + k_s x = 0.$$

From Newton's second law, the second-order differential equation motion is

$$m\frac{d^2x}{dt^2} + k_s x = 0.$$

The Hamiltonian function is expressed as

$$H(x, p) = \ + \ = \frac{1}{2}(mv^2 \quad k_s x^2) = \frac{1}{2} \ \frac{1}{m} p^2 \quad k_s x^2 \div .$$

From the Hamiltonian equations of motion $\dot{p}_i = \frac{H}{q_i}$ and $\dot{q}_i = \frac{H}{p_i}$, as given by (2.3.4), one obtains

$$\dot{p} = \frac{H}{x} = \ k_s x,$$

$$\dot{x} = \dot{q} = \frac{H}{p} = \frac{p}{m}.$$

The equivalence the results and equations of motion are obvious.

2.4. ATOMIC STRUCTURES AND QUANTUM MECHANICS

The fundamental and applied research as well as engineering developments in NEMS and MEMS have undergone major developments in last years. High-performance nanostructures and nanodevices, as well as MEMS have been manufactured and implemented (accelerometers and microphones, actuators and sensors, molecular wires and transistors, et cetera). Smart structures and MEMS have been mainly designed and built using conventional electromechanical and CMOS technologies. The next critical step to be made is to research nanoelectromechanical structures and systems, and these developments will have a tremendous positive impact on economy and society. Nanoengineering studies NEMS and MEMS, as well as their structures and subsystems, which are made from atoms and molecules, and the electron is considered as a fundamental particle. The students and engineers have obtained the necessary background in physics classes. The properties and performance of materials (media) is understood through the analysis of the atomic structure.

The atomic structures were studied by Rutherford and Einstein (in the 1900's), Heisenberg and Dirac (in the 1920's), Schrödinger, Bohr, Feynman, and many other scientists. For example, the theory of quantum electrodynamics studies the interaction of electrons and photons. In the 1940's, the major breakthrough appears in augmentation of the electron dynamics with electromagnetic field. One can control molecules and group of molecules (nanostructures) applying the electromagnetic field, and micro- and nanoscale devices (e.g., actuators and sensors) have been fabricated, and some problems in structural design and optimization have been approached and solved. However, these nano- and micro-scale devices (which have dimensions nano- and micrometers) must be controlled, and one faces an extremely challenging problem to design NEMS and MEMS integrating control and optimization, self-organization and decision making, diagnostics and self-repairing, signal processing and communication, as well as other features. In 1959, Richard Feynman gave a talk to the American Physical Society in which he emphasized the important role of nanotechnology and nanoscale organic and inorganic systems on the society and progress.

All media are made from atoms, and the medium properties depend on the atomic structure. Recalling the Rutherford's structure of the atomic nuclei, we can view here very simple atomic model and omit detailed composition, because only three subatomic particles (proton, neutron and electron) have bearing on chemical behavior.

The nucleus of the atom bears the major mass. It is an extremely dense region, which contains positively charged protons and neutral neutrons. It occupies small amount of the atomic volume compared with the virtually indistinct cloud of negatively charged electrons attracted to the positively charged nucleus by the force that exists between the particles of opposite electric charge.

For the atom of the element the number of protons is always the same but the number of neutrons may vary. Atoms of a given element, which differ in number of neutrons (and consequently in mass), are called isotopes. For example, carbon always has 6 protons, but it may have 6 neutrons as well. In this case it is called “carbon-12” (^{12}C). The representation of the carbon atom is given in Figure 2.4.1.

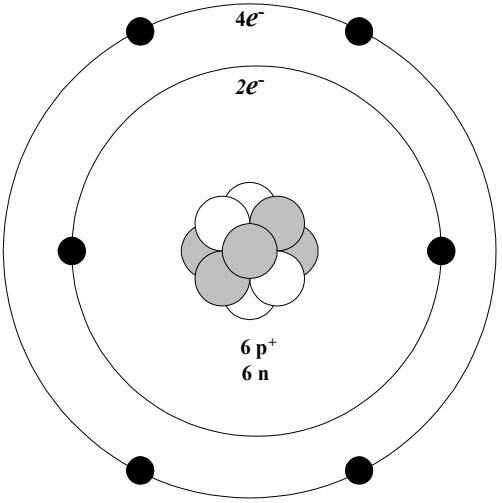


Figure 2.4.1.Simplified two-dimensional representation of carbon atom (C). Six protons (p+, dashed color) and six neutrons (n, white) are in centrally located nucleus. Six electrons (e^-, black), orbiting the nucleus, occupy two shells

Atom has no net charge due to the equal number of positively charged protons in the nucleus and negatively charged electrons around it. For example, all atoms of carbon have 6 protons and 6 electrons. If electrons are lost or gained by the neutral atom due to the chemical reaction, a charged particle called ion is formed.

When one deals with such subatomic particles as electron, the dual nature of matter places a fundamental limitation on how accurate we can describe both location and momentum of the object. Austrian physicist Erwin Schrödinger in 1926 derived an equation that describes wave and particle natures of the electron. This fundamental equation led to the new area in physics, called quantum mechanics, which enables us to deal with subatomic particles. The complete solution to Schrödinger’s equation gives a set of wave functions and set of corresponding energies. These wave functions are called orbitals. A collection of orbitals with the same principal quantum number, which describes the orbit, called electron shell. Each shell is divided into the number of subshells with the equal principal quantum

number. Each subshell consists of number of orbitals. Each shell may contain only two electrons of the opposite spin (Pouli exclusion principle). When the electron in the lowest energy orbital, the atom is in its ground state. When the electron enters the orbital, the atom is in an excited state. To promote the electron to the excited-state orbital, the photon of the appropriate energy should be absorbed as the energy supplement.

When the size of the orbital increases, and the electron spends more time farther from the nucleus. It possesses more energy and less tightly bound to the nucleus. The most outer shell is called the valence shell. The electrons, which occupy it, are referred as valence electrons. Inner shells electrons are called the core electrons. There are valence electrons, which participate in the bond formation between atoms when molecules are formed, and in ion formation when the electrons are removed from the electrically neutral atom and the positively charged cation is formed. They possess the highest ionization energies (the energy which measure the easy of the removing the electron from the atom), and occupy energetically weakest orbital since it is the most remote orbital from the nucleus. The valence electrons removed from the valence shell become free electrons transferring the energy from one atom to another. We will describe the influence of the electromagnetic field on the atom later in the text, and it is relevant to include more detailed description of the Pauli exclusion principal.

The electric conductivity of a media is predetermined by the density of free electrons, and good conductors have the free electron density in the range of 10^{23} free electrons per cm^3. In contrast, the free electron density of good insulators is in the range of 10 free electrons per cm^3. The free electron density of semiconductors in the range from $10^7/cm^3$ to $10^{15}/cm^3$ (for example, the free electron concentration in silicon at 25^0C and 100^0C are $2\ \ 10^{10}/cm^3$ and $2\ \ 10^{12}/cm^3$, respectively). The free electron density is determined by the energy gap between valence and conduction (free) electrons. That is, the properties of the media (conductors, semiconductors, and insulators) are determined by the atomic structure.

Using the atoms as building blocks, one can manufacture different structures using the molecular nanotechnology. There are many challenging problems needed to be solve such as mathematical modeling and analysis, simulation and design, optimization and testing, implementation and deployment, technology transfer and mass production. In addition, to build NEMS, advanced manufacturing technologies must be developed and applied. To fabricate nanoscale systems at the molecular level, the problems in atomic-scale positional assembly (“maneuvering things atom by atom" as Richard Feynman predicted) and artificial self-replication (systems are able to build copies of themselves, e.g., like the crystals growth process, complex DNA strands which copy tens of millions atoms with perfect accuracy, or self replicating tomato which has millions of genes, proteins, and other molecular components) must be solved. The author does not encourage the blind copying, and the submarine and whale are very different even though both sail. Using the Scanning or Atomic Probe Microscopes, it is possible to

achieve positional accuracy in the angstrom-range. However, the atomic-scale "manipulator" (which will have a wide range of motion guaranteeing the flexible assembly of molecular components), controlled by the external source (electromagnetic field, pressure, or temperature) must be designed and used. The position control will be achieved by the molecular computer and which will be based on molecular computational devices.

The quantitative explanation, analysis and simulation of natural phenomena can be approached using comprehensive mathematical models which map essential features. The Newton laws and Lagrange equations of motion, Hamilton concept and d'Alambert concept allow one to model conventional mechanical systems, and the Maxwell equations applied to model electromagnetic phenomena. In the 1920's, new theoretical developments, concepts and formulations (*quantum mechanics*) have been made to develop the atomic scale theory because atomic-scale systems do not obey the classical laws of physics and mechanics. In 1900 Max Plank discovered the effect of quantization of energy, and he found that the radiated (emitted) energy is given as

$E = nhv$,

where n is the nonnegative integer, $n = 0, 1, 2, \ldots$; h is the Plank constant, $h = 6.626 \quad 10^{-34}$ J-sec; v is the frequency of radiation, $v = \frac{c}{l}$, c is the speed of light, $c = 3 \quad 10^{8} \frac{\text{m}}{\text{sec}}$; l is the wavelength which is measured in angstroms ($\overset{o}{A} = 1 \quad 10^{-10}$ m), $l = \frac{c}{v}$.

The following discrete energy values result:

$E_0 = 0$, $E_1 = hv$, $E_2 = 2hv$, $E_3 = 3hv$, etc.

The observation of discrete energy spectra suggests that each particle has the energy hv (the radiation results due to N particles), and the particle with the energy hv is called *photon*.

The *photon* has the momentum as expressed as

$$p = \frac{hv}{c} = \frac{h}{l}.$$

Soon, Einstein demonstrated the discrete nature of light, and Niels Bohr develop the model of the hydrogen atom using the planetary system analog, see Figure 2.4.2. It is clear that if the electron has planetary-type orbits, it can be excited to an outer orbit and can "fall" to the inner orbits. Therefore, to develop the model, Bohr postulated that the electron has the certain stable circular orbit (that is, the orbiting electron does not produces the radiation because otherwise the electron would lost the energy and change the path); the electron changes the orbit of higher or lower energy by receiving or radiating discrete amount of energy; the angular momentum of the electron is $p = nh$.

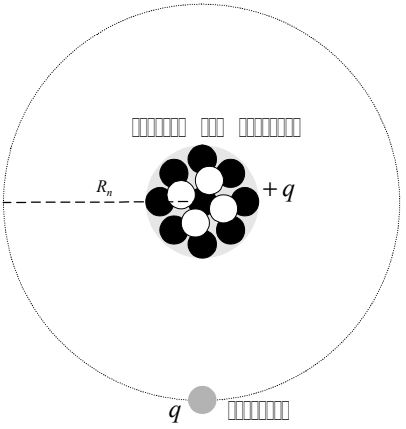


Figure 2.4.2. Hydrogen atom: uniform circular motion

To attain the uniform circular motion, using Newton's law, the electrostatic (Coulomb) force must be equal to the radial force, and for radii R_1 and R_2 we have

$$\frac{q^2}{4\pi\varepsilon_0 R_1^2} = \frac{mv^2}{R_1} \text{ and } \frac{q^2}{4\pi\varepsilon_0 R_2^2} = \frac{mv^2}{R_2}.$$

That is, in general

$$\frac{q^2}{4\pi\varepsilon_0 R_n^2} = \frac{mv^2}{R_n},$$

where R_n is the radius of the n orbit, and $R_n = \dfrac{4\pi\varepsilon_0 n^2 h^2}{mq^2}$.

Applying the expression for the angular momentum

$p = nh = mvR_n$,

we have

$$\frac{q^2}{4\pi\varepsilon_0 R_n^2} = \frac{m}{R_n}\left(\frac{nh}{mR_n}\right)^2 = \frac{1}{mR_n}\frac{n^2h^2}{R_n^2}.$$

The kinetic and potential energies are

$$= \tfrac{1}{2}mv^2 = \frac{mq^4}{32\pi^2\varepsilon_0^2 n^2 h^2} \text{ and } = \frac{q^2}{4\pi\varepsilon_0 R_n} = \frac{mq^4}{16\pi^2\varepsilon_0^2 n^2 h^2}.$$

The total energy of the electron in the nth orbit is found to be

$$E_n = + = \frac{mq^4}{32\pi^2\varepsilon_0^2 n^2 h^2}.$$

One finds the energy difference between the orbits as

$$E = E_{n2} \quad E_{n1} = \frac{mq^4}{32p^2 e_0^2 h^2} \quad \frac{1}{n_1^2} \quad \frac{1}{n_2^2} \div .$$

Bohr's model was expanded and generalized by Heisenberg and Schrödinger using the *matrix* and *wave mechanics*. The characteristics of particles and waves are augmented replacing the trajectory consideration by the waves using continuous, finite, and single-valued wave function

(x, y, z, t) in the Cartesian coordinate system,

(r, f, z, t) in the cylindrical coordinate system,

(r, q, f, t) in the spherical coordinate system.

The wavefunction gives the dependence of the wave amplitude on space coordinates and time.

Using the classical mechanics, for a particle of mass *m* with energy *E* moving in the Cartesian coordinate system one has

$$\underset{\textit{total energy}}{E(x, y, z, t)} = \underset{\textit{kinetic energy}}{(x, y, z, t)} + \underset{\textit{potential energy}}{(x, y, z, t)}$$

$$= \frac{p^2(x, y, z, t)}{2m} + \quad (x, y, z, t) = \underset{\textit{Hamiltonian}}{H(x, y, z, t)}.$$

Thus, we have

$$p^2(x, y, z, t) = 2m\left[E(x, y, z, t) \qquad (x, y, z, t)\right].$$

Using the formula for the wavelength (Broglie's equation)

$$l = \frac{h}{p} = \frac{h}{mv},$$

one finds

$$\frac{1}{l^2} = \left(\frac{p}{h}\right)^2 \div = \frac{2m}{h^2}\left[E(x, y, z, t) \qquad (x, y, z, t)\right].$$

This expression is substituted in the *Helmholtz* equation

$$^2 \quad + \frac{4p^2}{l^2} \quad = 0$$

which gives the evolution of the wavefunction.

We obtain the Schrödinger equation as

$$E(x, y, z, t) \quad (x, y, z, t) = \quad \frac{\hbar^2}{2m} \quad ^2 \quad (x, y, z, t) + \quad (x, y, z, t) \quad (x, y, z, t)$$

or

$$E(x, y, z, t) \quad (x, y, z, t)$$

$$= \quad \frac{\hbar^2}{2m} \quad \frac{^2 \quad (x, y, z, t)}{x^2} + \frac{^2 \quad (x, y, z, t)}{y^2} + \frac{^2 \quad (x, y, z, t)}{z^2} \div$$

$$+ \quad (x, y, z, t) \quad (x, y, z, t).$$

Here, the modified Plank constant is

$\hbar = \frac{h}{2p} = 1.055 \quad 10^{-34}$ J-sec.

In 1926, Erwine Schrödinger derive the following equation

$$\frac{\hbar^2}{2m} \quad ^2 \quad + \quad = E$$

which can be related to the Hamiltonian

$$H = \quad \frac{\hbar^2}{2m} \quad + \quad ,$$

and thus

$$H \quad = E \quad .$$

For different coordinate systems we have

Cartesian system

$$^2 \quad (x, y, z, t)$$

$$= \frac{^2 \quad (x, y, z, t)}{x^2} + \frac{^2 \quad (x, y, z, t)}{y^2} + \frac{^2 \quad (x, y, z, t)}{z^2};$$

cylindrical system

$$^2 \quad (r, f, z, t)$$

$$= \frac{1}{r} \frac{}{r} \quad r \frac{(r, f, z, t)}{r} \div + \frac{1}{r^2} \frac{^2 \quad (r, f, z, t)}{f^2} + \frac{^2 \quad (r, f, z, t)}{z^2};$$

spherical system

$$^2 \quad (r, q, f, t) =$$

$$\frac{1}{r^2} \frac{}{r} \quad r^2 \frac{(r, q, f, t)}{r} \div + \frac{1}{r^2 \sin q} \frac{}{q} \quad \sin q \frac{(r, q, f, t)}{q} \div$$

$$+ \frac{1}{r^2 \sin^2 q} \frac{^2 \quad (r, q, f, t)}{f^2}.$$

The Schrödinger partial differential equation must be solved, and the wavefunction is normalized using the probability density

$$\left| \quad \right|^2 dV = 1.$$

Let us illustrate the application of the Schrödinger equation.

Example 2.4.1.

Assume that the particle moves in the *x* direction (translational motion). We have,

$$\frac{\hbar^2}{2m} \frac{d^2 \quad (x)}{dx^2} + \quad (x) \quad (x) = E(x) \quad (x).$$

The Hamiltonian function is given as

$$H(x,p)=\frac{p^2(x)}{2m}+\quad(x)=\quad\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+\quad(x).$$

Let the particle moves from $x = 0$ to $x = x_f$, and the potential energy is

$$(x)=\begin{array}{l}0,\qquad 0\quad x\quad x_f\\ ,\ x<0\ \text{and}\ x>x_f\end{array}.$$

Thus, the motion of the particle is bounded in the "potential wall", and

$$(x)=\begin{array}{l}\text{continuous if}\ 0\quad x\quad x_f\\ 0\ \text{if}\ x<0\ \text{and}\ x>x_f\end{array}.$$

If $0\quad x\quad x_f$, the potential energy is zero, and we have

$$\frac{\hbar^2}{2m}\frac{d^2\quad(x)}{dx^2}=E\quad(x),\ 0\quad x\quad x_f.$$

The solution of the resulting second-order differential equation

$$\frac{d^2\quad(x)}{dx^2}+k^2\quad(x)=0,\ k=\sqrt{\frac{2mE}{\hbar^2}}$$

is

$$(x)=ae^{ikx}+be^{\ ikx}=a(\cos kx+i\sin kx)+b(\cos kx\quad i\sin kx)$$

$$=c\sin kx+d\cos kx.$$

The solution can be easily verified by plugging the solution in the left-side of the differential equation

$$\frac{\hbar^2}{2m}\frac{d^2\quad(x)}{dx^2}=E\quad(x),$$

and we have

$$E\quad(x)=E\quad(x).$$

It should be emphasized that the kinetic energy of the particle is given as $\frac{p^2}{2m}$, where $p = kh$.

It is obvious that one must use the boundary conditions.

We have $(x)\big|_{x=0}=\quad(0)=0$, and therefore $d = 0$.

From $(x)\big|_{x=x_f}=\quad(x_f)=0$ using $c\sin kx_f=0$ one must find the constant c and the expression for kx_f.

Assuming that $c\quad 0$ from $c\sin kx_f=0$, we have

$$kx_f=np\ ,$$

where n is the positive or negative integer (if $n = 0$, the wavefunction vanishes everywhere, and thus, $n\quad 0$).

From $k = \sqrt{\frac{2mE}{\hbar^2}}$ and making use of $kx_f = np$ we have the expression for the energy (discrete values of the energy which allow of solution of the Schrödinger equation) as

$$E_n = \frac{\hbar^2 p^2}{2mx_f^2} n^2, n = 1, 2, 3, \ldots,$$

where the integer *n* designates the allowed energy level (*n* is called the quantum number).

For example, if *n* = 1 and *n* = 2, we have $E_1 = \frac{\hbar^2 p^2}{2mc^2}$ (the lowest possible energy which is called the ground state) and $E_2 = \frac{2\hbar^2 p^2}{mc^2}$.

Thus, we have illustrated that the energy of the particle is quantized.

The expression for the wavefunction is found to be

$$_n(x) = c \sin kx + d \cos kx = c \sin \frac{np}{x_f} x .$$

Using the probability density, we normalize the wavefunction, and the following results

$$\int_0^{x_f} {}^2_n(x)dx = c^2 \int_0^{x_f} \sin^2 \frac{np}{x_f} x dx = c^2 \frac{x_f}{np} \int_0^{np} \sin^2 g dg$$

$$c^2 \frac{x_f}{np} \frac{np}{2} = c^2 \frac{x_f}{2} = 1, g = \frac{np}{x_f} x.$$

Hence, $c = \sqrt{\frac{2}{x_f}}$, and one finally obtains

$$_n(x) = \sqrt{\frac{2}{x_f}} \sin \frac{np}{x_f} x, \quad 0 \quad x \quad x_f .$$

For *n* = 1 and *n* = 2, we have

$$_1(x) = \sqrt{\frac{2}{x_f}} \sin \frac{p}{x_f} x \text{ and } \quad _2(x) = \sqrt{\frac{2}{x_f}} \sin \frac{2p}{x_f} x .$$

Using the formula for the probability density, as given by $r = \quad ^T$, one has

$$r_n(x) = \frac{2}{x_f} \sin^2 \frac{np}{x_f} x .$$

It was shown that

$H = E$, $H = \frac{\hbar^2}{2m} +$.

Using the CGS (centimeter/gram/second) units, when the electromagnetic field is quantized, the potential can be used instead of wavefunction. In particular, using the momentum operator due to electron orbital angular momentum **L**, the classical Hamiltonian for electrons in electromagnetic field is

$$H = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 ef .$$

From the Hamilton equations

$$\dot{q} = \frac{H}{p} \text{ and } \dot{p} = -\frac{H}{q},$$

by making use of

$$\frac{d\mathbf{r}}{dt} = \frac{1}{m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right), \mathbf{p} = m\mathbf{v} \quad \frac{e}{c}\mathbf{A} , \dot{\mathbf{p}} = \frac{e}{mc}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)\times\frac{\mathbf{A}}{\mathbf{x}} + e\frac{f}{\mathbf{x}},$$

one finds the Lorentz force equation

$$\mathbf{F} = \frac{e}{c}\mathbf{v} \quad \mathbf{B} \quad e\mathbf{E} .$$

This equation gives the force due to motion in a magnetic field and the force due to electric field.

It is important to emphasize that the following equation results

$$\frac{\hbar^2}{2m} \quad ^2 + \frac{e}{2mc}\mathbf{B}\times\mathbf{L} \quad + \frac{e^2}{8mc^2}\left(r^2B^2 \quad (\mathbf{r}\times\mathbf{B})^2\right) \quad = \left(E + ef\right)$$

to study the quantized Hamilton equation, where the dominant term due to magnetic field is

$$\frac{e}{2mc}\mathbf{B}\times\mathbf{L} = \text{ì} \times\mathbf{B},$$

where **ì** the magnetic momentum due to the electron orbital angular momentum (the so-called Zeeman effect) is **ì** $= \frac{e}{2mc}\mathbf{L}$.

2.5. MOLECULAR AND NANOSTRUCTURE DYNAMICS

Conventional, mini- and microscale electromechanical systems can be modeled using electromagnetic and circuitry theories, classical mechanics and thermodynamic, as well as other fundamental concepts. The complexity of mathematical models of mini- and microelectromechanical systems (nonlinear ordinary and partial differential equations explicitly describe the spectrum of electromagnetics and electromechanics phenomena and processes) is not ambiguous, and numerical algorithms to solve the equations derived are available. Illustrated examples have been studied in sections 2.2 and 2.3. Nano-scale structures, in general, cannot be studied using the conventional concepts, and the basis of quantum mechanics was covered in chapter 2.4.

The fundamental and applied research in molecular nanotechnology and nanostructures, nanodevices and nanosystems, NEMS and MEMS, is concentrated on design, modeling, simulation, and fabrication of molecular scale structures and devices. The design, modeling, and simulation of NEMS, MEMS, and their components can be attacked using advanced theoretical developments and simulation concepts. Comprehensive analysis must be performed before the designer embarks in costly fabrication (a wide range of nano-scale structures and devices, molecular machines and subsystems, can be fabricated with atomic precision) because through modeling and simulation the rapid evaluation and prototyping can be performed facilitating significant advantages and manageable perspectives to attain the desired objectives. With advanced computer-aided-design tools, complex large-scale nanostructures, nanodevices, and nanosystems can be designed, analyzed, and evaluated.

Classical quantum mechanics does not allow the designer to perform analytical and numerical analysis even for simple nanostructures which consist of a couple of molecules. Steady-state three-dimensional modeling and simulation are also restricted to simple nanostructures. Our goal is to develop a fundamental understanding of phenomena and processes in nanostructures with emphasis on their further applications in nanodevices, nanosubsystems, NEMS, and MEMS. The objective is the development of theoretical fundamentals (theory of nanoelectromechanics) to perform 3D+ (three-dimensional geometry dynamics in time domain) modeling and simulation.

The atomic level electomechanics can be studied using the wave function solving the Schrödinger equation for N-electron systems (multi-body problem). However, this problem cannot be solved even for simple nanostrustures. In papers [2 - 4], the density functional theory was developed, and the charge density is used rather than the electron wavefunctions. In particular, the N-electron problem is formulated as N one-electron equations where each electron interacts with all other electrons via an effective exchange-correlation potential. These interactions are

augmented using the charge density. Plane wave sets and total energy pseudo-potential methods can be used to solve the Kohn-Sham one electron equations [2 - 4]. The Hellmann-Feynman theory can be applied to calculate the forces solving the molecular dynamics problem [1 - 5].

2.5.1. Schrödinger Equation and Wavefunction Theory

For two point charges, Coulomb's law is given as

$$\mathbf{F} = \frac{q_1 q_2}{4\pi\varepsilon d^2}\mathbf{a}_r = \frac{q_1 q_2}{4\pi\varepsilon}\frac{(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3},$$

and in the Cartesian coordinate systems one has

$$\mathbf{F} = \frac{q_1 q_2}{4\pi\varepsilon d^2}\mathbf{a}_r = \frac{q_1 q_2}{4\pi\varepsilon d^2}\frac{(x-x')\mathbf{a}_x + (y-y')\mathbf{a}_y + (z-z')\mathbf{a}_z}{\sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}}.$$

In the case of charge distribution, using the volume charge density ρ_v, the net force exerted on q_1 by the entire volume charge distribution is the vector sum of the contribution from all differential elements of charge within this distribution. In particular,

$$\mathbf{F} = \frac{q_1}{4\pi\varepsilon}\int_v \rho_v \frac{(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|^3}dv,$$

see Figure 2.5.1.

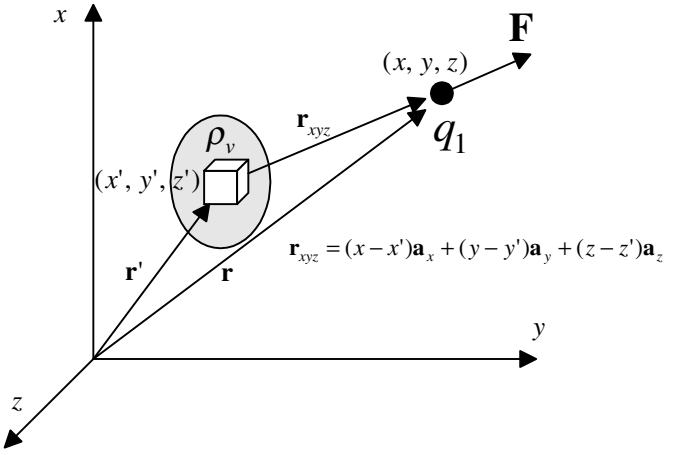


Figure 2.5.1. Coulomb's law

In the electrostatic field, the potential energy stored in a region of continuous charge distribution is found as

$$\Pi_V = \tfrac{1}{2}\int_v \mathbf{D}\cdot\mathbf{E}dv = \tfrac{1}{2}\int_v \varepsilon\mathbf{E}^2 dv = \tfrac{1}{2}\int_v \rho_v(\mathbf{r})V(\mathbf{r})dv\,,$$

where $V(\mathbf{r})$ is the potential; v is the volume containing ρ_v.

The charge distribution can be given in terms of volume, surface, and line charges. In particular, we have

$$V(\mathbf{r}) = \int_v \frac{\rho_v(\mathbf{r}')}{4\pi\varepsilon\left|\mathbf{r}-\mathbf{r}'\right|}dv'\,,$$

$$V(\mathbf{r}) = \int_s \frac{\rho_s(\mathbf{r}')}{4\pi\varepsilon\left|\mathbf{r}-\mathbf{r}'\right|}ds'\,,$$

and $V(\mathbf{r}) = \int_l \frac{\rho_l(\mathbf{r}')}{4\pi\varepsilon\left|\mathbf{r}-\mathbf{r}'\right|}dl'$.

It should be emphasized that that the electric field intensity is found as

$$\mathbf{E}(\mathbf{r}) = \int_v \frac{\rho_v(\mathbf{r}')}{4\pi\varepsilon}\frac{(\mathbf{r}-\mathbf{r}')}{\left|\mathbf{r}-\mathbf{r}'\right|^3}dv'\,.$$

Thus, the energy of an electric field or a charge distribution is stored in the field.

The energy, stored in the steady magnetic field is

$$\Pi_M = \tfrac{1}{2}\int_v \mathbf{B}\cdot\mathbf{H}dv\,.$$

The Hamiltonian function, which in section 2.4 was given as

$$H = \underbrace{-\frac{\hbar^2}{2m}\nabla^2}_{\text{one-electron kinetic energy}} + \underbrace{\Pi}_{\text{potential energy}}\,,$$

was used to derive the one-electron Schrödinger equation.

To describe the behavior of electrons in a media, one must use N-dimensional Schrödinger equation to obtain the N-electron wavefunction $\Psi\left(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N\right)$.

The Hamiltonian for an isolated *N*-electron atomic system is

$$H = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 - \frac{\hbar^2}{2M}\nabla^2 - \sum_{i=1}^{N}\frac{1}{4\pi\varepsilon}\frac{e_i q}{\left|\mathbf{r}_i-\mathbf{r}_n^{'}\right|} + \sum_{i\neq j}^{N}\frac{1}{4\pi\varepsilon}\frac{e^2}{\left|\mathbf{r}_i-\mathbf{r}_j^{'}\right|}\,,$$

where q is the potential due to nucleus; $e = 1.6\times10^{-19}$ C.

For an isolated *N*-electron, *Z*-nucleus molecular system, the Hamiltonian function (Hamiltonian operator) is found to be

$$H = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 - \sum_{k=1}^{Z}\frac{\hbar^2}{2m_k}\nabla_k^2$$
$$-\sum_{i=1}^{N}\sum_{k=1}^{Z}\frac{1}{4\pi\varepsilon}\frac{e_i q_k}{\left|\mathbf{r}_i - \mathbf{r}_k^{'}\right|} + \sum_{i\neq j}^{N}\frac{1}{4\pi\varepsilon}\frac{e^2}{\left|\mathbf{r}_i - \mathbf{r}_j^{'}\right|} + \sum_{k\neq m}^{Z}\frac{1}{4\pi\varepsilon}\frac{q_k q_m}{\left|\mathbf{r}_k - \mathbf{r}_m^{'}\right|},$$

where q_k are the potentials due to nuclei.

Terms of the Hamiltonian function $-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2$ and $-\sum_{k=1}^{Z}\frac{\hbar^2}{2m_k}\nabla_k^2$ are the multi-body kinetic energy operators.

Term $-\sum_{i=1}^{N}\sum_{k=1}^{Z}\frac{1}{4\pi\varepsilon}\frac{e_i q_k}{\left|\mathbf{r}_i - \mathbf{r}_k^{'}\right|}$ maps the interaction of the electrons with the nuclei at **R** (the electron-nucleus attraction energy operator).

In the Hamiltonian, the fourth term $\sum_{i\neq j}^{N}\frac{1}{4\pi\varepsilon}\frac{e^2}{\left|\mathbf{r}_i - \mathbf{r}_j^{'}\right|}$ gives the interactions of electrons with each other (the electron-electron repulsion energy operator).

Term $\sum_{k\neq m}^{Z}\frac{1}{4\pi\varepsilon}\frac{q_k q_m}{\left|\mathbf{r}_k - \mathbf{r}_m^{'}\right|}$ describes the interaction of the Z nuclei at **R** (the nucleus-nucleus repulsion energy operator).

For an isolated *N*-electron *Z*-nucleus atomic or molecular systems in the Born-Oppenheimer nonrelativistic approximation, we have

$H\Psi = E\Psi$.

Thus, the Schrödinger equation is

$$\left[-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 - \sum_{k=1}^{Z}\frac{\hbar^2}{2m_k}\nabla_k^2\right.$$
$$\left.-\sum_{i=1}^{N}\sum_{k=1}^{Z}\frac{1}{4\pi\varepsilon}\frac{e_i q_k}{\left|\mathbf{r}_i - \mathbf{r}_k^{'}\right|} + \sum_{i\neq j}^{N}\frac{1}{4\pi\varepsilon}\frac{e^2}{\left|\mathbf{r}_i - \mathbf{r}_j^{'}\right|} + \sum_{k\neq m}^{Z}\frac{1}{4\pi\varepsilon}\frac{q_k q_m}{\left|\mathbf{r}_k - \mathbf{r}_m^{'}\right|}\right]$$
$$\times\Psi\left(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N\right) = E\left(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N\right)\Psi\left(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N\right).$$

(2.5.1)

The total energy $E\left(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N\right)$ must be found using the nucleus-nucleus Coulomb repulsion energy as well as the electron energy.

It is very difficult, or impossible, to solve analytically or numerically the nonlinear partial differential equation (2.5.1). Taking into account only the Coulomb force (electrons and nuclei are assumed to interact due to the Coulomb force only), the Hartree approximation is applied. In particular, the

N-electron wavefunction $\Psi(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N)$ is expressed as a product of N one-electron wavefunctions as

$$\Psi(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{N-1},\mathbf{r}_N)=\psi_1(t,\mathbf{r}_1)\psi_2(t,\mathbf{r}_2)...\psi_{N-1}(t,\mathbf{r}_{N-1})\psi_N(t,\mathbf{r}_N).$$

The one-electron Schrödinger equation for *j*th electron is

$$\left(-\frac{\hbar^2}{2m}\nabla_j^2+\Pi(t,\mathbf{r})\right)\psi_j(t,\mathbf{r})=E_j(t,\mathbf{r})\psi_j(t,\mathbf{r}). \tag{2.5.2}$$

In equation (2.5.2), the first term $-\frac{\hbar^2}{2m}\nabla_j^2$ is the one-electron kinetic energy, and $\Pi(t,\mathbf{r}_j)$ is the total potential energy. The potential energy includes the potential that *j*th electron feels from the nucleus (considering the ion, the repulsive potential in the case of anion, or attractive in the case of cation). It is obvious that *j*th electron feels the repulsion (repulsive forces) from other electrons. Assumed that the negative electrons charge density $\rho(\mathbf{r})$ is smoothly distributed in **R**. Hence, the potential energy due interaction (repulsion) of an electron in **R** is

$$\Pi_{Ej}(t,\mathbf{r})=\int_{\mathbf{R}}\frac{e\rho(\mathbf{r}')}{4\pi\varepsilon|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'.$$

We made some assumptions, and the results derived contradict with some fundamental principles. The Pauli exclusion principle requires that the multi-system wavefunction is an antisymmetric under the interchange of electrons. For two electrons, we have,

$$\Psi(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_j,...,\mathbf{r}_{j+i},...,\mathbf{r}_{N-1},\mathbf{r}_N)=-\Psi(t,\mathbf{r}_1,\mathbf{r}_2,...,\mathbf{r}_{j+i},...,\mathbf{r}_j,...,\mathbf{r}_{N-1},\mathbf{r}_N).$$

This principle is not satisfied, and the generalizations is needed to integrate the asymmetry phenomenon using the asymmetric coefficient ± 1. The Hartree-Fock equation is

$$\left[-\frac{\hbar^2}{2m}\nabla_j^2+\Pi(t,\mathbf{r})\right]\psi_j(t,\mathbf{r}) - \sum_i\int_{\mathbf{R}}\frac{\psi_i^*(t,\mathbf{r}')\psi_j(t,\mathbf{r}')\psi_i(t,\mathbf{r})\psi_j^*(t,\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'=E_j(t,\mathbf{r})\psi_j(t,\mathbf{r}). \tag{2.5.3}$$

The so-called Hartree-Fock nonlinear partial differential equation (2.5.3), which is difficult to solve, is the approximation because the multi-body electron interactions should be considered in general. Thus, the explicit equation for the total energy must be used. This phenomenon can be integrated using the charge density function.

2.5.2. Density Functional Theory

There is a critical need to develop computationally efficient and accurate procedures to perform quantum modeling of nano-scale structures. This section reports the related results and gives the formulation of the modeling problem to avoid the complexity associated with many-electron wavefunctions which result if the classical quantum mechanics formulation is used. The complexity of the Schrödinger equation is enormous even for very simple molecules. For example, the carbon atom has 6 electrons. Can one visualize six-dimensional space? Furthermore, the simplest carbon nanotube molecule has 6 carbon atoms. That is, one has 36 electrons, and 36-dimensional problem results. The difficulties associated with the solution of the Schrödinger equation drastically limit the applicability of the conventional quantum mechanics. The analysis of properties, processes, phenomena, and effects in even simplest nanostructures cannot be studied and comprehended. The problems can be solved applying the Hohenberg-Kohn density functional theory.

The statistical consideration, proposed by Thomas and Fermi in 1927, gives the distribution of electrons in atoms. The following assumptions were used: electrons are distributed uniformly, and there is an effective potential field that is determined by the nuclei charge and the distribution of electrons. Considering electrons distributed in a three-dimensional box, the energy analysis can be performed. Summing all energy levels, one finds the energy. Thus, one can relate the total kinetic energy and the electron charge density. The statistical consideration can be used in order to approximate the distribution of electrons in an atom. The relation between the total kinetic energy of N electrons E, and the electron density was derived using the local density approximation concept. The Thomas-Fermi kinetic energy functional is

$$\Gamma_F\left(\rho_e(\mathbf{r})\right)=2.87\int_{\mathbf{R}}\rho_e^{5/3}(\mathbf{r})d\mathbf{r},$$

and the exchange energy is found to be

$$E_F\left(\rho_e(\mathbf{r})\right)=0.739\int_{\mathbf{R}}\rho_e^{4/3}(\mathbf{r})d\mathbf{r}.$$

For homogeneous atomic systems, the application of the electron charge density $\rho_e(\mathbf{r})$, considering electrostatic electron-nucleus attraction and electron-electron repulsion, Thomas and Fermi derived the following energy functional

$$E_F\left(\rho_e(\mathbf{r})\right)=2.87\int_{\mathbf{R}}\rho_e^{5/3}(\mathbf{r})d\mathbf{r}-q\int_{\mathbf{R}}\frac{\rho_e(\mathbf{r})}{r}d\mathbf{r}+\int_{\mathbf{R}}\int_{\mathbf{R}}\frac{1}{4\pi\varepsilon}\frac{\rho_e(\mathbf{r})\rho_e(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}'.$$

Following this idea, instead of the many-electron wavefunctions, Kohn proposed to use the charge density for N-electron systems [2, 4]. Only the knowledge of the charge density is needed to perform analysis of molecular dynamics. The charge density is the function that describes the number of

electrons per unit volume (function of three spatial variables *x*, *y* and *z* in the Cartesian coordinate system). The quantum mechanics and quantum modeling must be applied to understand and analyze nanostructures and nanodevices because they operate under the quantum effects.

The total energy of *N*-electron system under the external field is defined in the term of the three-dimensional charge density $\rho(\mathbf{r})$ [1 - 5]. The complexity is significantly decreased because the problem of modeling of *N*-electron *Z*-nucleus systems become equivalent to the solution of equation for one electron. The total energy is given as

$$E\left(t,\rho(\mathbf{r})\right)=\underbrace{\Gamma_1\left(t,\rho(\mathbf{r})\right)}_{\text{kinetic energy}}+\Gamma_2\left(t,\rho(\mathbf{r})\right)+\underbrace{\int_{\mathbf{R}}\frac{e\rho(\mathbf{r}')}{4\pi\varepsilon|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'}_{\text{potential energy}}, \qquad (2.5.4)$$

where $\Gamma_1\left(t,\rho(\mathbf{r})\right)$ and $\Gamma_2\left(t,\rho(\mathbf{r})\right)$ are the interacting (exchange) and non-interacting kinetic energies of a single electron in *N*-electron *Z*-nucleus system,

$$\Gamma_1\left(t,\rho(\mathbf{r})\right)=\int_{\mathbf{R}}\gamma\left(t,\rho(\mathbf{r})\right)\rho(\mathbf{r})d\mathbf{r}\,,\ \Gamma_2\left(t,\rho(\mathbf{r})\right)=-\frac{\hbar^2}{2m}\sum_{j=1}^{N}\int_{\mathbf{R}}\psi_j^*(t,\mathbf{r})\nabla_j^2\psi_j(t,\mathbf{r})d\mathbf{r}\,;$$

$\gamma\left(t,\rho(\mathbf{r})\right)$ is the parameterization function.

It should be emphasized that the Kohn-Sham electronic orbitals are subject to the following orthogonal condition

$$\int_{\mathbf{R}}\psi_i^*(t,\mathbf{r})\psi_j(t,\mathbf{r})d\mathbf{r}=\delta_{ij}\,.$$

The state of substance (media) depends largely on the balance between the kinetic energies of the particles and the interparticle energies of attraction.

The expression for the total potential energy is easily justified.

Term $\int_{\mathbf{R}}\frac{e\rho(\mathbf{r}')}{4\pi\varepsilon|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'$ represents the Coulomb interaction in **R**, and the total potential energy is a functions of the charge density $\rho(\mathbf{r})$.

The total kinetic energy (interactions of electrons and nuclei, and electrons) is integrated into the equation for the total energy. The total energy, as given by (2.5.4), is stationary with respect to variations in the charge density. The charge density is found taking note of the Schrödinger equation. The first-order Fock-Dirac electron charge density matrix is

$$\rho_e(\mathbf{r})=\sum_{j=1}^{N}\psi_j^*(t,\mathbf{r})\psi_j(t,\mathbf{r})\,. \qquad (2.5.5)$$

The three-dimensional electron charge density is a function in three variables (*x*, *y* and *z* in the Cartesian coordinate system). Integrating the electron charge density $\rho_e(\mathbf{r})$, one obtains the charge of the total number of electrons *N*. Thus,

$$\int_{\mathbf{R}} \rho_e(\mathbf{r}) d\mathbf{r} = Ne \, .$$

Hence, $\rho_e(\mathbf{r})$ satisfies the following properties

$$\rho_e(\mathbf{r}) > 0 \, , \int_{\mathbf{R}} \rho_e(\mathbf{r}) d\mathbf{r} = Ne \, ,$$

$$\int_{\mathbf{R}} \left| \sqrt{\nabla \rho_e(\mathbf{r})} \right|^2 d\mathbf{r} < \infty \, ,$$

$$\int_{\mathbf{R}} \nabla^2 \rho_e(\mathbf{r}) d\mathbf{r} = \infty \, .$$

For the nuclei charge density, we have

$$\rho_n(\mathbf{r}) > 0 \text{ and } \int_{\mathbf{R}} \rho_n(\mathbf{r}) d\mathbf{r} = \sum_{k=1}^{Z} q_k \, .$$

There exist an infinite number of antisymmetric wavefunctions that give the same $\rho(\mathbf{r})$. The minimum-energy concept (energy-functional minimum principle) is applied. The total energy is a function of $\rho(\mathbf{r})$, and the so-called ground state Ψ must minimize the expectation value $\langle E(\rho) \rangle$.

The searching density functional $F(\rho)$, which searches all Ψ in the N-electron Hilbert space *H* to find $\rho(\mathbf{r})$ and guarantee the minimum to the energy expectation value, is expressed as

$$F(\rho) \le \min_{\substack{\Psi \to \rho \\ \Psi \in H_\Psi}} \langle \Psi | \, E(\rho) \, | \Psi \rangle \, ,$$

where H_Ψ is any subset of the N-electron Hilbert space.

Using the variational principle, we have

$$\frac{\Delta E(\rho)}{\Delta f(\rho)} = \int_{\mathbf{R}} \frac{\Delta E(\rho)}{\Delta \rho(\mathbf{r}')} \frac{\Delta \rho(\mathbf{r}')}{\Delta f(\mathbf{r})} d\mathbf{r}' = 0 \, ,$$

where $f(\rho)$ is the nonnegative function.

Thus, $\left. \dfrac{\Delta E(\rho)}{\Delta f(\rho)} \right|_N = \text{const} \, .$

The solutions to the system of equations (2.5.2) is found using the charge density (2.5.5).

To perform the analysis of nanostructure dynamics, one studies the molecular dynamics. The force and displacement must be found. Substituting the expression for the total kinetic and potential energies in (2.5.4), where the charge density is given by (2.5.5), the total energy $E\big(t, \rho(\mathbf{r})\big)$ results.

The external energy is supplied to control nanoscale actuators, and one has

$$E_{\Sigma}(t,\mathbf{r}) = E_{external}(t,\mathbf{r}) + E(t,\rho(\mathbf{r})).$$

Then, the force at position $\mathbf{r}_r$ is

$$\mathbf{F}_r(t,\mathbf{r}) = -\frac{dE_{\Sigma}(t,\mathbf{r})}{d\mathbf{r}_r}$$

$$= -\frac{\partial E_{\Sigma}(t,\mathbf{r})}{\partial \mathbf{r}_r} - \sum_j \frac{\partial E(t,\mathbf{r})}{\partial \psi_j(t,\mathbf{r})}\frac{\partial \psi_j(t,\mathbf{r})}{\partial \mathbf{r}_r} - \sum_j \frac{\partial E(t,\mathbf{r})}{\partial \psi_j^*(t,\mathbf{r})}\frac{\partial \psi_j^*(t,\mathbf{r})}{\partial \mathbf{r}_r}. \quad (2.5.6)$$

Taking note of

$$\sum_j \frac{\partial E(t,\mathbf{r})}{\partial \psi_j(t,\mathbf{r})}\frac{\partial \psi_j(t,\mathbf{r})}{\partial \mathbf{r}_r} + \sum_j \frac{\partial E(t,\mathbf{r})}{\partial \psi_j^*(t,\mathbf{r})}\frac{\partial \psi_j^*(t,\mathbf{r})}{\partial \mathbf{r}_r} = 0,$$

the expression for the force is found from (2.5.6). In particular, one finds

$$\mathbf{F}_r(t,\mathbf{r}) = -\frac{\partial E_{external}(t,\mathbf{r})}{\partial \mathbf{r}_r}$$

$$-\int_{\mathbf{R}} \rho(t,\mathbf{r})\frac{\partial[\Pi_r(t,\mathbf{r}) + \Gamma_r(t,\mathbf{r})]}{\partial \mathbf{r}_r} d\mathbf{r} - \int_{\mathbf{R}} \frac{\partial E_{\Sigma}(t,\mathbf{r})}{\partial \rho(t,\mathbf{r})}\frac{\partial \rho(t,\mathbf{r})}{\partial \mathbf{r}_r} d\mathbf{r}.$$

As the wavefunctions converge (the conditions of the Hellmann-Feynman theorem are satisfied), we have

$$\int_{\mathbf{R}} \frac{\partial E(t,\mathbf{r})}{\partial \rho(t,\mathbf{r})}\frac{\partial \rho(t,\mathbf{r})}{\partial \mathbf{r}_r} d\mathbf{r} = 0.$$

One can deduce the expression for the wavefunctions, find the charge density, calculate the forces, and study processes and phenomena in nanoscale. The displacement is found using the following equation of motion

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}_r(t,\mathbf{r}),$$

or

$$m\frac{d^2(\vec{x},\vec{y},\vec{z})}{dt^2} = \mathbf{F}_r(\vec{x},\vec{y},\vec{z}).$$

2.5.3. Nanostructures and Molecular Dynamics

Atomistic modeling can be performed using the force field method. The effective interatomic potential for a system of N particles is found as the sum of the second-, third-, fourth-, and higher-order terms as

$$\Pi(\mathbf{r}_1,...,\mathbf{r}_N)=\sum_{i,j=1}^{N}\Pi^{(2)}(\mathbf{r}_{ij})+\sum_{i,j,k=1}^{N}\Pi^{(3)}(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k)+\sum_{i,j,k,l=1}^{N}\Pi^{(4)}(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k,\mathbf{r}_l)+...$$

Usually, the interatomic effective pair potential $\sum_{i,j=1}^{N}\Pi^{(2)}(\mathbf{r}_{ij})$, which depends on the interatomic distance r_{ij} between the nuclei i and j, dominates. For example, the three-body interconnection terms cannot be omitted only if the angle-dependent potentials are considered. Using the effective ionic charges Q_i and Q_j, we have

$$\Pi^{(2)}=\underbrace{\frac{Q_iQ_j}{4\pi\varepsilon r_{ij}}}_{electrostatic}+\underbrace{\phi(\mathbf{r}_{ij})}_{short-range},$$

where $\phi(\mathbf{r}_{ij})$ is the short-range interaction energy due to the repulsion between electron charge clouds, Van der Waals attraction, bond bending and stretching phenomena.

For ionic and partially ionic media we have

$$\phi(r_{ij})=k_{1ij}e^{-k_{2ij}r_{ij}}-k_{3ij}r_{ij}^{-6}+k_{4ij}r_{ij}^{-12},$$

where $k_{1ij}=\sqrt{k_{1i}k_{1j}}$, $k_{2ij}=\sqrt{k_{2i}k_{2j}}$, $k_{3ij}=\sqrt{k_{3i}k_{3j}}$ and $k_{4ij}=\sqrt{k_{4i}k_{4j}}$; k_i are the bond energy constants (for example, for *Si* we have Q = 2.4, k_3 = 0.00069 and k_4 = 104, for *Al* one has Q = 1.4, k_3 = 1690 and k_4 = 278, and for Na^+ we have Q = 1, k_3 = 0.00046 and k_4 = 67423).

Another, commonly used approximation is $\phi(r_{ij})=k_{5ij}(r_{ij}-r_{Eij})$, where r_{ij} is the bond length, $r_{ij}=|\mathbf{r}_j-\mathbf{r}_i|$; r_{Eij} is the equilibrium bond distance

Performing the summations in the studied **R**, one finds the potential energy, and the force results. The position (displacement) is represented by the vector **r** which in the Cartesian coordinate system has the components *x*, *y* and *z*. Taking note of the expression for the potential energy $\Pi(\vec{\mathbf{r}})=\Pi(\mathbf{r}_1,...,\mathbf{r}_N)$, one has

$$\sum\vec{F}(\vec{\mathbf{r}})=-\nabla\Pi(\vec{\mathbf{r}}).$$

From Newton's second law for the system of N particles, we have the following equation of motion

$$m_N\frac{d^2\vec{\mathbf{r}}_N}{dt^2}+\nabla\Pi(\vec{\mathbf{r}}_N)=0,$$

or

$$m_i \frac{d^2(\vec{x}_i, \vec{y}_i, \vec{z}_i)}{dt^2} + \frac{\partial \Pi(\vec{x}_i, \vec{y}_i, \vec{z}_i)}{\partial(\vec{x}_i, \vec{y}_i, \vec{z}_i)} = 0, i = 1,..., N .$$

To perform molecular modeling one applies the energy-based methods. It was shown that electrons can be considered explicitly. However, it can be assumed that electrons will obey the optimum distribution once the positions of the nuclei in **R** are known. This assumption is based on the Born-Oppenheimer approximation of the Schrödinger equation. This approximation is satisfied because nuclei mass is much greater then electron mass, and thus, nuclei motions (vibrations and rotations) are slow compared with the electrons' motions. Therefore, nuclei motions can be studied separately from electrons dynamics. Molecules can be studied as *Z*-body systems of elementary masses (nuclei) with springs (bonds between nuclei). The molecule potential energy (potential energy equation) is found using the number of nuclei and bond types (bending, stretching, lengths, geometry, angles, and other parameters), van der Waals radius, parameters of media, etc. The molecule potential energy surface is

$$E_T = E_{bs} + E_b + E_{sb} + E_{ts} + E_W + E_{dd} .$$

Here, the energy due to bond stretching is found using the equation similar to Hook's law. In particular,

$$E_{bs} = k_{bs1}(l - l_0) + k_{bs3}(l - l_0)^3 ,$$

where k_{bs1} and k_{bs3} are the constants; l and l_0 are the actual and natural bond length (displacement).

The equations for energies due to bond angle bending E_b, stretch-bend interactions E_{sb}, torsion strain E_{ts}, van der Waals interactions E_W, and dipole-dipole interactions E_{dd} are well known and can be readily applied.

2.6. MOLECULAR WIRES AND MOLECULAR CIRCUITS

The molecular wire consists of the single molecule chain with its end adsorbed to the surface of the gold lead that can cover monolayers of other molecules. Molecular wires connect the nanoscale structures and devices. The current density of carbon nanotubes, 1,4-dithiol benzene (molecular wire) and copper are 10^{11}, 10^{12} and 10^6 electroncs/sec-nm^2, respectively. The current technology allows one to fill carbon nanotubes with other media (metals, organic and inorganic materials). That is, to connect nanostructures, as shown in Figure 2.6.1, it is feasible to use molecular wires which can be synthesized through the organic synthesis.

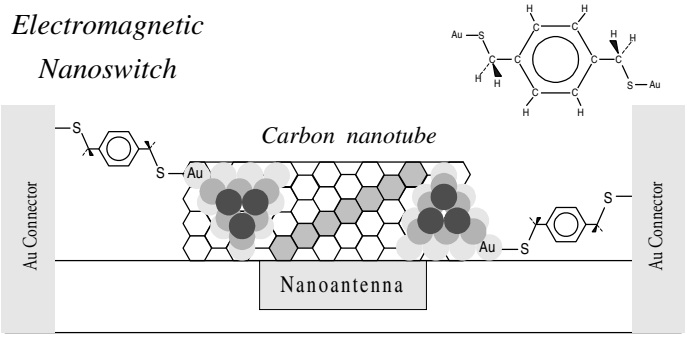


Figure 2.6.1. Nanoswitch with carbon nanotube, molecular wire (1,4-dithiol benzene) and nanoantenna

Consider covalent bonds. These bonds occur from sharing the electrons between two atoms. Covalent bonds represent the interactions of two nonmetallic elements, or metallic and nonmetallic elements. Let us study the electron density around the nuclei of two atoms. If electron clouds overlap region passes through on the line joining two nuclei, the bond is called σ bond, see Figure 2.6.2. The overlap may occur between orbitals perpendicularly oriented to the internuclear axis. The resulting covalent bond produces overlap above and below the internuclear axis. Such bond is called π bond. There is no probability of finding the electron on the internuclear axis in a π bond, and the overlap in it is lesser than in the σ bond. Therefore, π bonds are generally weaker than σ bonds.

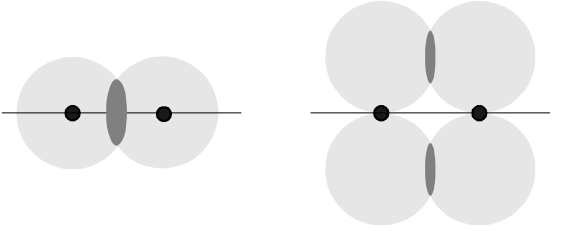


Figure 2.6.2. σ and π covalent bonds

Single bonds are usually σ bonds. Double bonds, which are much stronger, consist of one σ bond and one π bond, and the triple bond (the strongest one) consists of one σ bond and two π bonds. In the case of carbon nanotubes, the strong interaction among the carbon atoms is guaranteed by the strength of the C-C single bond which holds carbon atoms together in the honeycomb-like hexagon unit (open-ended nanotube).

In molecular wires, the current i_m is a function of the applied voltage u_m, and Landauer's formula is

$$i_m = \frac{2e}{h}\int_{-\infty}^{+\infty} T(E_m, u_m)\left(\frac{1}{e^{\frac{E_m-\mu_{p1}}{k_BT}}+1} - \frac{1}{e^{\frac{E_m-\mu_{p2}}{k_BT}}+1}\right)dE_m ,$$

where μ_{p1} and μ_{p2} are the electrochemical potentials, $\mu_{p1} = E_F + \frac{1}{2}eu_m$ and $\mu_{p2} = E_F - \frac{1}{2}eu_m$; E_F is the equilibrium Fermi energy of the source; $T(E_m, u_m)$ is the transmission function obtained using the molecular energy levels and coupling.

We have [7]

$$i_m = \frac{2e}{h}\int_{\mu_{p1}}^{\mu_{p2}} T(E_m, u_m)\frac{1}{4k_BT}\operatorname{sech}^2\left(\frac{E_m}{2k_BT}\right)dE_m , \; k_BT = 26\text{meV}.$$

Thus, the molecular wire conductance is found as

$$c_m = \frac{\partial i_m}{\partial u_m} \approx \frac{e^2}{h}\left[T(\mu_{p1}) + T(\mu_{p2})\right].$$

Using molecular wires and molecular circuits (which form molecular electronic switches and devices), the designer can synthesize polyphenylene-based rectifying diodes, switching logics, as well as other devices. It must be emphasized that the results given above are based upon the thorough and comprehensive overview of molecular circuits reported in [2]. Figure 2.6.3 illustrates the molecular circuitry for a polyphenylene-based molecular rectifying diode. This diode can be fabricated using the chemically doped polyphenylene-based molecular wire as the constructive medium. The electron donating substituent group *X* (n-dopant) and the electron withdrawing substituent group *Y* (p-dopant) form two intermolecular dopant groups. These groups are separated by the semi-insulating group *R* (potential energy barrier) from an electron acceptor subcomplex. Thus, the *R* group serves as an insulation (barrier) between the donor *X* and acceptor *Y*. The semi-insulating group *R* can be synthesized using the aliphatic (sigma-bounded methylene) or dimethylene groups. To guarantee electrical isolation between the molecular circuitry and gold substrate, additional barrier is used as shown in Figure 2.6.3.

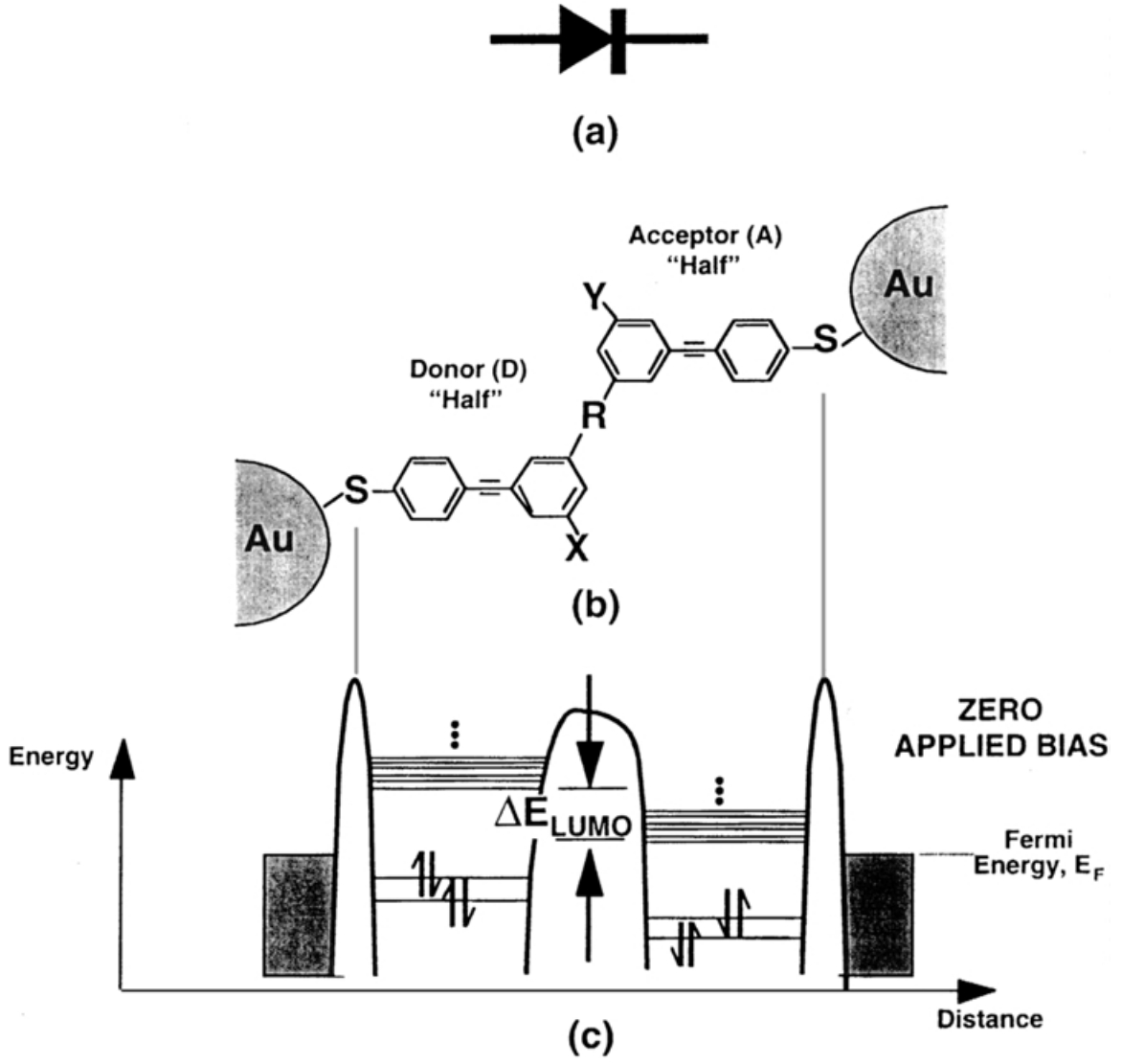
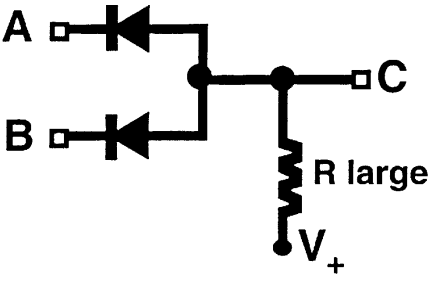
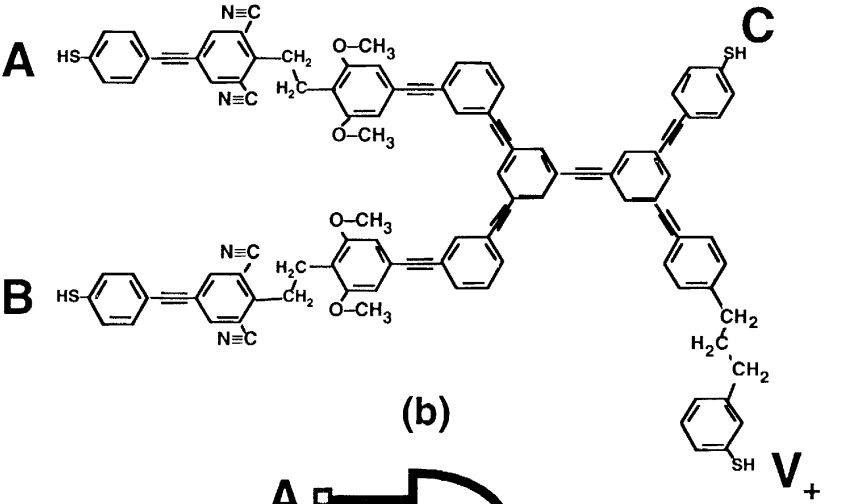


Figure 2.6.3. Molecular circuit and schematics of electron orbital energylevels levels for a polyphenylene-based molecular rectifying diode [2]

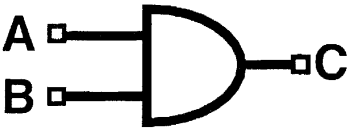
In computers, DSPs, microcontrollers, and microprocessors, simple arithmetic functions, e.g. addition and subtraction, are implemented using combinational register-level components. Adders and subtracters (which have carry-in and carry-out lines) of fixed-point binary numbers are basic register-level components from which other arithmetic circuits are formed. Other arithmetic components are widely used, and comparators compare the magnitude of two binary numbers. These arithmetic elements can be fabricated using molecular circuit technology. In fact, to perform logic operations (AND, OR, XOR, and NOT gates) and arithmetic, diode-based molecular electronic digital circuits and nanologic gates can be synthesized using single nanoscale molecule structures. It should be emphasized that the size of these molecular logic gates is within 5 nm (thousand times less then the logic gates used in current computers which are fabricated using most advanced CMOS technologies). Using diode-diode logic, AND and OR logic gates are designed using molecular circuits, and the schematics are illustrated in Figures 2.6.4 and 2.6.5. The molecular AND logic gate is designed by connecting in parallel two diodes. The doped polyphenylene-based diodes are connected through polyphenylene-based wire. The semi-insulating group *R* (potential energy barrier) reduces power dissipation and maintains a distinct output voltage signal at the terminal *C* when the *A* and *B* inputs (carry-in lines) cause the molecular diodes to be forward biased (current flows through diodes). The difference between the AND and OR gates is that the diode orientations, see Figures 2.6.4 and 2.6.5. The diode-based molecular electronic digital circuit (XOR gate) is illustrated in Figure 2.6.6, and the truth table is also documented. The total voltage applied across the XOR gate is the sum of the voltage drop across the input resistances plus the voltage drop across the resonant tunneling diode (RTD). The effective resistance of the logic gate, containing two rectifying diodes, differs whether one or both parallel signals (*A* and *B* can be 1 or 0) are *on*. If *A* and *B* are *on* (1), the effective resistance is half. Thus, according to Ohm's law, there are two possible cases: full voltage drop and half voltage drop which distinct the XOR gate operating points. Figure 2.6.7 documents the molecular half adder which is synthesized using the AND and XOR molecular gates. Here, *A* and *B* denote the one-bit binary signals (inputs) to the adder, while *S* (sum bit) and *C* (carry bit) are one-bit binary signals (outputs). The XOR gate gives the sum of two bits, and the resulting output is at lead *S*. The AND gate forms the sum of two bits, and the resulting output is at lead *C*. The molecular full adder is given in Figure 2.6.8.



(a)



(b)

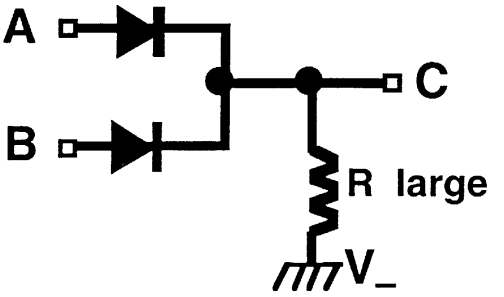


(c)

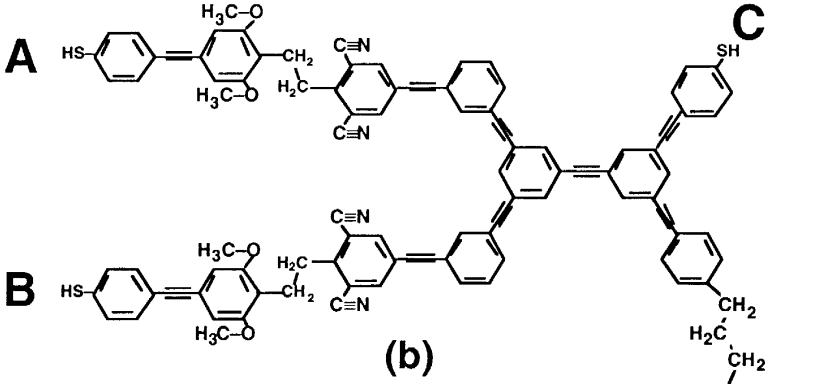
A	B	C
0	0	0
1	0	0
0	1	0
1	1	1

(d)

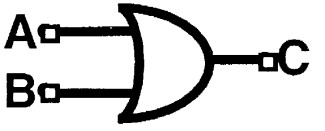
Figure 2.6.4. Molecular circuit and schematics of AND molecular gate [2]



(a)



(b)

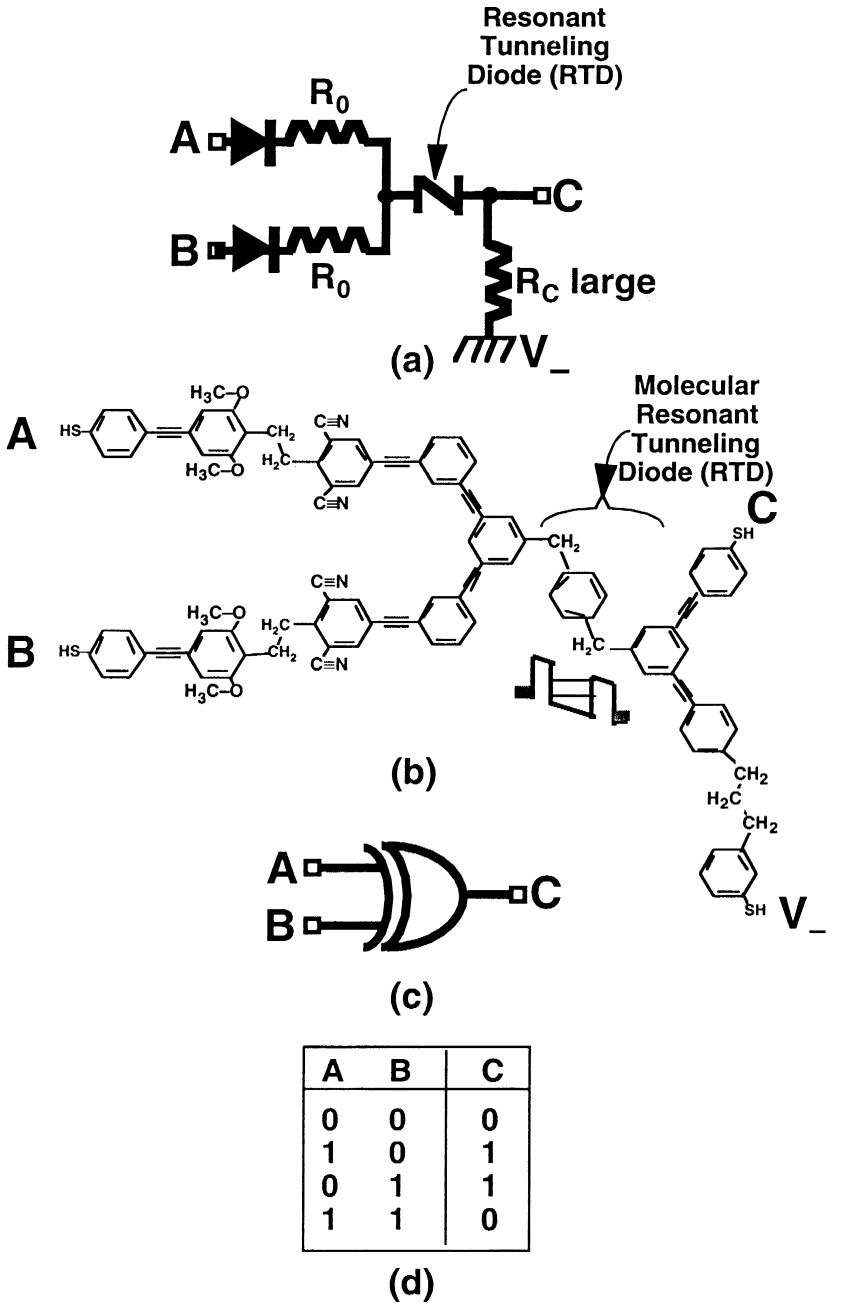


(c)

A	B	C
0	0	0
1	0	1
0	1	1
1	1	1

(d)

Figure 2.6.5. Molecular circuit and schematics of OR molecular gate [2]



A	B	C
0	0	0
1	0	1
0	1	1
1	1	0

(d)

Figure 2.6.6. Molecular circuit and schematics of XOR molecular gate [2]

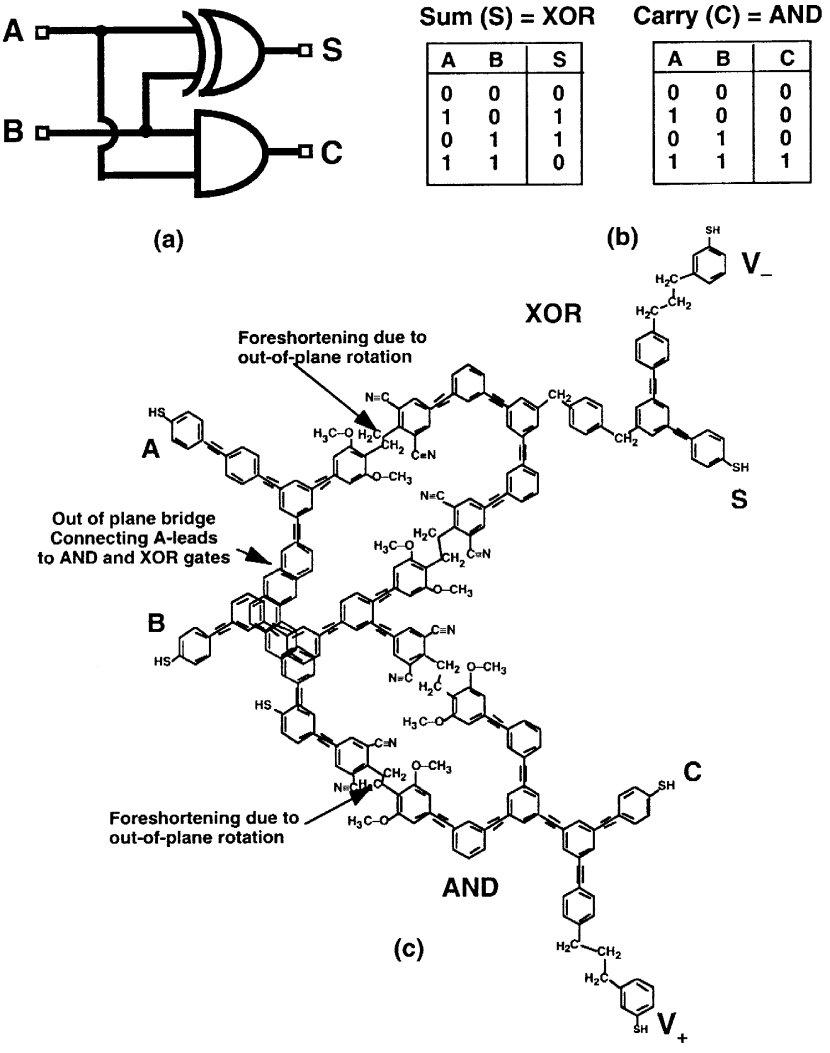


Figure 2.6.7. Molecular circuit and schematics of molecular half adder [2]

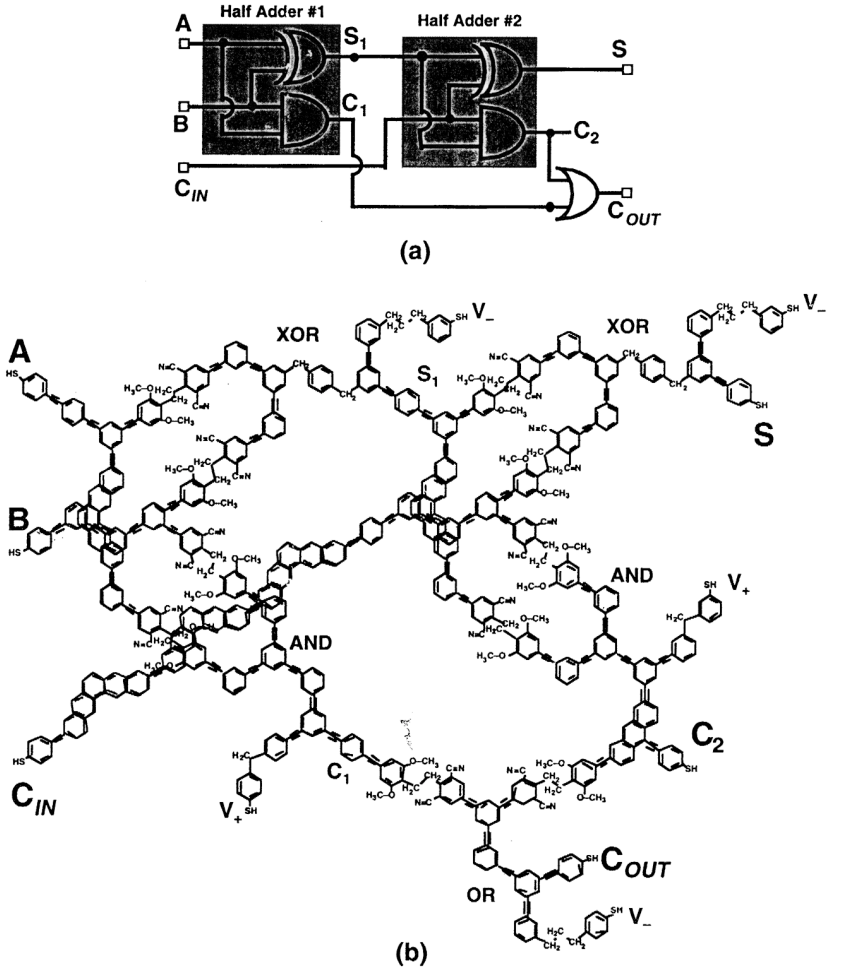


Figure 2.6.8. Molecular circuit and schematics of molecular full adder [2]

References

1. E. R. Davidson, *Reduced Density Matrices in Quantum Chemistry,* Academic Press, New York, NY, 1976.
2. J. C. Ellenbogen and J. C. Love, *Architectures for molecular electronic computers*, MP 98W0000183, MITRE Corporation, 1999.
3. P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev*., vol. 136, pp. B864-B871, 1964.
4. W. Kohn and R. M. Driezler, "Time-dependent density-fuctional theory: conceptual and practical aspects," *Phys. Rev. Letters*, vol. 56, pp. 1993 - 1995, 1986.
5. W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev*., vol. 140, pp. A1133 - A1138, 1965.
6. R. G. Parr and W Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, NY, 1989.
7. W. T. Tian, S. Datta, S. Hong, R. Reifenberger, J. I. Henderson, and C. P. Kubiak, "Conductance spectra of molecular wires," *Int. Journal Chemical Phisics*, vol. 109, no. 7, pp. 2874-2882, 1998.

2.7. THERMOANALYSIS AND HEAT EQUATION

It is known that the heat propagates (flows) in the direction of decreasing temperature, and the rate of propagation is proportional to the gradient of the temperature. Using the thermal conductivity of the media k_t and the temperature $T(t,x,y,z)$, one has the following equation to calculate the velocity of the heat flow

$$\vec{\mathbf{v}}_h = -k_t \nabla T(t,x,y,z). \tag{2.7.1}$$

Consider the region **R** and let s is the boundary surface. Using the divergence theorem, from (2.7.1) one obtains the partial differential equation (heat equation) which is expressed as

$$\frac{\partial T(t,x,y,z)}{\partial t} = k^2 \nabla^2 T(t,x,y,z), \tag{2.7.2}$$

where k is the thermal diffusivity of the media.

We have

$$k = \frac{k_t}{k_h k_d},$$

where k_h and k_d are the specific heat and density constants.

Solving partial differential equation (2.7.2), which is subject to the initial and boundary conditions, one finds the temperature of the homogeneous media. In the Cartesian coordinate system, one has

$$\nabla^2 T(t,x,y,z) = \frac{\partial^2 T(t,x,y,z)}{\partial x^2} + \frac{\partial^2 T(t,x,y,z)}{\partial y^2} + \frac{\partial^2 T(t,x,y,z)}{\partial z^2}.$$

Using the Laplacian of T in the cylindrical and spherical coordinate systems, one can reformulate the thermoanalysis problem using different coordinates in order to straightforwardly solve the problem.

It the heat flow is steady (time-invariant), then

$$\frac{\partial T(t,x,y,z)}{\partial t} = 0.$$

Hence, three-dimensional heat equation (2.7.2) becomes Laplace's equation as given by

$$0 = k^2 \nabla^2 T(t,x,y,z).$$

The two-dimensional heat equation is

$$\frac{\partial T(t,x,y)}{\partial t} = k^2 \nabla^2 T(t,x,y) = k^2 \left(\frac{\partial^2 T(t,x,y)}{\partial x^2} + \frac{\partial^2 T(t,x,y)}{\partial y^2} \right).$$

If

$$\frac{\partial T(t,x,y)}{\partial t} = 0,$$

one has

$$0 = k^2 \nabla^2 T(t,x,y) = k^2 \left(\frac{\partial^2 T(t,x,y)}{\partial x^2} + \frac{\partial^2 T(t,x,y)}{\partial y^2} \right).$$

Using initial and boundary conditions, this partial differential equation can be solved using Fourier series, Fourier integrals, Fourier transforms.

The so-called one-dimensional heat equation is

$$\frac{\partial T(t,x)}{\partial t} = k^2 \frac{\partial^2 T(t,x)}{\partial x^2}$$

with initial and boundary conditions

$T(t_0,x) = T_t(x)$, $T(t,x_0) = T_0$ and $T(t,x_f) = T_f$.

A large number of analytical and numerical methods are available to solve the heat equation.

The analytic solution if

$T(t,x_0) = 0$ and $T(t,x_f) = 0$

is given as

$$T(t,x) = \sum_{i=1}^{\infty} B_i \sin \frac{i\pi x}{x_f} e^{-i^2 \frac{k^2 \pi^2}{x_f^2} t},$$

$$B_i = \frac{2}{x_f} \int_{x_0}^{x_f} T_t(x) \sin \frac{i\pi x}{x_f} dx.$$

Assuming that $T_t(x)$ is piecewise continuous in $x \in [x_0 \;\; x_f]$ and has one-sided derivatives at all interior points, one finds the coefficients of the Fourier sine series B_i.

Example 2.7.1.

Consider the copper bar with length 0.1 mm. The thermal conductivity, specific heat and density constants are $k_t = 1$, $k_h = 0.09$ and $k_d = 9$. The initial and boundary conditions are

$$T(0,x) = T_t(x) = 0.2 \sin \frac{\pi x}{0.001},\ T(t,0) = 0 \text{ and } T(t,0.001) = 0.$$

Find the temperature in the bar as a function of the position and time.

Solution.

From the general solution

$$T(t,x) = \sum_{i=1}^{\infty} B_i \sin \frac{i\pi x}{x_f} e^{-i^2 \frac{k^2 \pi^2}{x_f^2} t},$$

using the initial condition, we have

$$T(0,x) = \sum_{i=1}^{\infty} B_i \sin \frac{i\pi x}{x_f} = 0.2 \sin \frac{\pi x}{0.001}.$$

Thus, $B_1 = 0.2$ and all other B_i coefficients are zero.n

Hence, the solution (temperature as the function of the position and time) is found to be

$$T(t,x) = \sum_{i=1}^{\infty} B_i \sin\frac{i\pi x}{x_f} e^{-i^2\frac{k^2\pi^2}{x_f^2}t} = B_1 \sin\frac{\pi x}{x_f} e^{-\frac{k^2\pi^2}{x_f^2}t}$$

$$= 0.2\sin\frac{\pi x}{0.001} e^{-1.5\times10^7 t}.$$

CHAPTER 3

STRUCTURAL DESIGN, MODELING, AND SIMULATION

3.1. NANO- AND MICROELECTROMECHANICAL SYSTEMS

3.1.1. Carbon Nanotubes and Nanodevices

Carbon nanotubes, discovered in 1991, are molecular structures which consist of graphene cylinders closed at either end with caps containing pentagonal rings. Carbon nanotubes are produced by vaporizing carbon graphite with an electric arc under an inert atmosphere. The carbon molecules organize a perfect network of hexagonal graphite rolled up onto itself to form a hollow tube. Buckytubes are extremely strong and flexible and can be single- or multi-walled. The standard arc-evaporation method produces only multilayered tubes, and the single-layer uniform nanotubes (constant diameter) were synthesis only a couple years ago. One can fill nanotubes with any media, including biological molecules. The carbon nanotubes can be conducting or insulating medium depending upon their structure.

A single-walled carbon nanotube (one atom thick), which consists of carbon molecules, is illustrated in Figure 3.1.1. The application of these nanotubes, formed with a few carbon atoms in diameter, provides the possibility to fabricate devices on an atomic and molecular scale. The diameter of nanotube is 100000 times less that the diameter of the sawing needle. The carbon nanotubes, which are much stronger than steel wire, are the perfect conductor (better than silver), and have thermal conductivity better than diamond. The carbon nanotubes, manufactured using the carbon vapor technology, and carbon atoms bond together forming the pattern. Single-wall carbon nanotubes are manufactured using laser vaporization, arc technology, vapor growth, as well as other methods. Figure 3.1.2. illustrates the carbon ring with six atoms. When such a sheet rolls itself into a tube so that its edges join seamlessly together, a nanotube is formed.

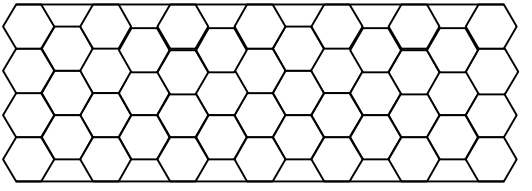


Figure 3.1.1. Single-walled carbon nanotube

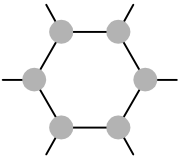


Figure 3.1.2. Single carbon nanotube ring with six atoms

Carbon nanotubes, which allow one to implement the molecular wire technology in nanoscale ICs, are used in NEMS and MEMS. Two slightly displaced (twisted) nanotube molecules, joined end to end, act as the diode. Molecular-scale transistors can be manufactured using different alignments. There are strong relationships between the nanotube electromagnetic properties and its diameter and degree of the molecule twist. In fact, the electromagnetic properties of the carbon nanotubes depend on the molecule's twist, and Figures 3.1.3 illustrate possible configurations. If the graphite sheet forming the single-wall carbon nanotube is rolled up perfectly (all its hexagons line up along the molecules axis), the nanotube is a perfect conductor. If the graphite sheet rolls up at a twisted angle, the nanotube exhibits the semiconductor properties. The carbon nanotubes, which are much stronger than steel wire, can be added to the plastic to make the conductive composite materials.

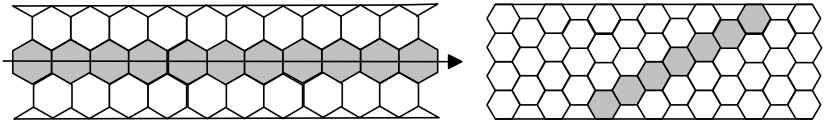


Figure 3.1.3. Carbon nanotubes

The vapor grown carbon nanotubes with N layers are illustrated in Figure 3.1.4, and the industrially manufactured nanotubes have Δngstroms diameter and length.

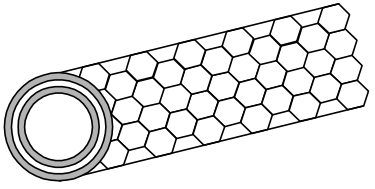
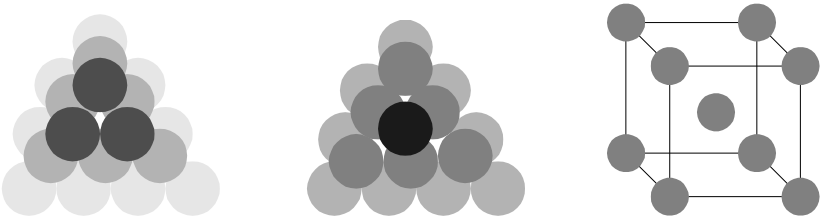


Figure 3.1.4. N-layer carbon nanotube

The carbon nanotubes can be organized as large-scale complex neural networks to perform computing and data storage, sensing and actuation, etc. The density of ICs designed and manufactured using the carbon nanotube technology thousands time exceed the density of ICs developed using convention silicon and silicon-carbide technologies.

Metallic solids (conductor, for example copper, silver, and iron) consist of metal atoms. These metallic solids usually have hexagonal, cubic, or body-centered cubic close-packed structures (see Figure 3.1.5). Each atom has 8 or 12 adjacent atoms. The bonding is due to valence electrons that are delocalized thought the entire solid. The mobility of electrons is examined to study the conductivity properties.



(a) (b) (c)

Figure 3.1.5. Close packing of metal atoms: a) cubic packing; b) hexagonal packing; c) body-centered cubic

More than two electrons can fit in an orbital. Furthermore, these two electrons must have two opposite spin states (spin-up and spin-down). Therefore, the spins are said to be paired. Two opposite directions in which the electron spins (up $+\frac{1}{2}$ and down $-\frac{1}{2}$) produce oppositely directed magnetic fields. For an atom with two electrons, the spin may be either parallel (S = 1) or opposed and thus cancel (S = 0). Because of spin pairing, most molecules have no net magnetic field, and these molecules are called *diamagnetic* (in the absence of the external magnetic field, the net magnetic field produced by the magnetic fields of the orbiting electrons and the magnetic fields produced by the electron spins is zero). The external magnetic field will produce no torque on the *diamagnetic* atom as well as no realignment of the dipole fields. Accurate quantitative analysis can be performed using the quantum theory. Using the simplest atomic model, we assume that a positive nucleus is surrounded by electrons which orbit in various circular orbits (an electron on the orbit can be studied as a current loop, and the direction of current is opposite to the direction of the electron rotation). The torque tends to align the magnetic field, produced by the orbiting electron, with the external magnetic field. The electron can have a spin magnetic moment of $\pm 9 \ \ 10^{-24}$ A-m^2. The plus and minus signs that there are two possible electron alignments; in particular, aiding or opposing to the external magnetic field. The atom has many electrons, and only the spins of those electrons in shells which are not completely filed contribute to the atom magnetic moment. The nuclear spin negligible contributes to the atom moment. The magnetic properties of the media (diamagnetic, paramagnetic, superparamagnetic, ferromagnetic, antiferromagnetic, ferrimagnetic) result due to the combination of the listed atom moments

Let us discuss the *paramagnetic* materials. The atom can have small magnetic moment, however, the random orientation of the atoms results that the net torque is zero. Thus, the media do not show the magnetic effect in the absence the external magnetic field. As the external magnetic field is applied, due to the atom moments, the atoms will align with the external field. If the atom has large dipole moment (due to electron spin moments), the material is called ferromagnetic. In antiferromagnetic materials, the net magnetic moment is zero, and thus the ferromagnetic media are only slightly affected by the external magnetic field.

Using carbon nanotubes, one can design electromechanical and electromagnetic nanoswitches, which are illustrated in Figure 3.1.6.

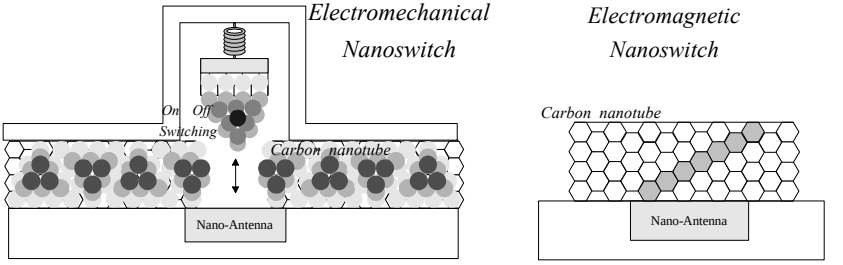


Figure 3.1.6. Application of carbon nanotubes in nanoswitches

3.1.2. Microelectromechanical Systems and Microdevices

Different MEMS have been discussed, and it was emphasized that MEMS can be used as actuators, sensors, and actuators-sensors. Due to the limited torque and force densities, MEMS usually cannot develop high torque and force, and large-scale cooperative MEMS are used, e.g. multilayer configurations. In contrast, these characteristics (power, torque, and force densities) are not critical in sensor applications. Therefore, MEMS are widely used as sensors. Signal-level signals, measured by sensors, are fed to analog or digital controllers, and sensor design, signal processing, and interfacing are extremely important in engineering practice. Smart integrated sensors are the sensors in which in addition to sensing the physical variable, data acquisition, filtering, data storage, communication, interfacing, and networking are embedded. Thus, while the primary component is the sensing element (microstructure), multifunctional integration of sensors and ICs is the current demand. High-performance accelerometers, manufactured by Analog Devices using integrated microelectromechanical system technology (*i*MEMS), are studied in this section. In addition, the application of smart integrated sensors is discussed.

We study the dual-axis, surface-micromachined ADXL202 accelerometer (manufactured on a single monolithic silicon chip) which combines highly accurate acceleration sensing motion microstructure (proof mass) and signal processing electronics (signal conditioning ICs). As documented in the Analog Device Catalog data (which is attached), this accelerometer, which is manufactured using the *i*MEMS technology, can measure dynamic positive and negative acceleration (vibration) as well as static acceleration (force of gravity). The functional block diagram of the ADXL202 accelerometer with two digital outputs (ratio of pulse width to period is proportional to the acceleration) is illustrated in Figure 3.1.7.

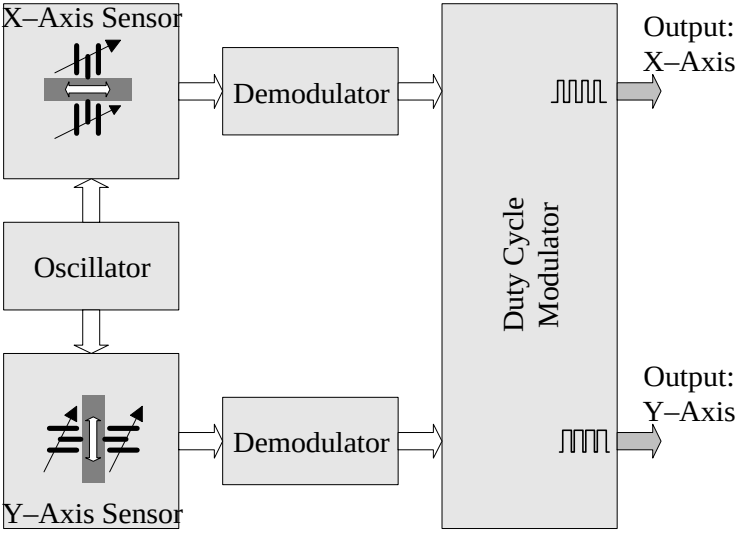


Figure 3.1.7. Functional block diagram of the ADXL202 accelerometer

Polysilicon surface-micromachined sensor motion microstructure is fabricated on the silicon wafer by depositing polysilicon on the sacrificial oxide layer which is then etched away leaving the suspended proof mass (beam). Polysilicon springs suspend this proof mass over the surface of the wafer. The deflection of the proof mass is measured using the capacitance difference, see Figure 3.1.8.

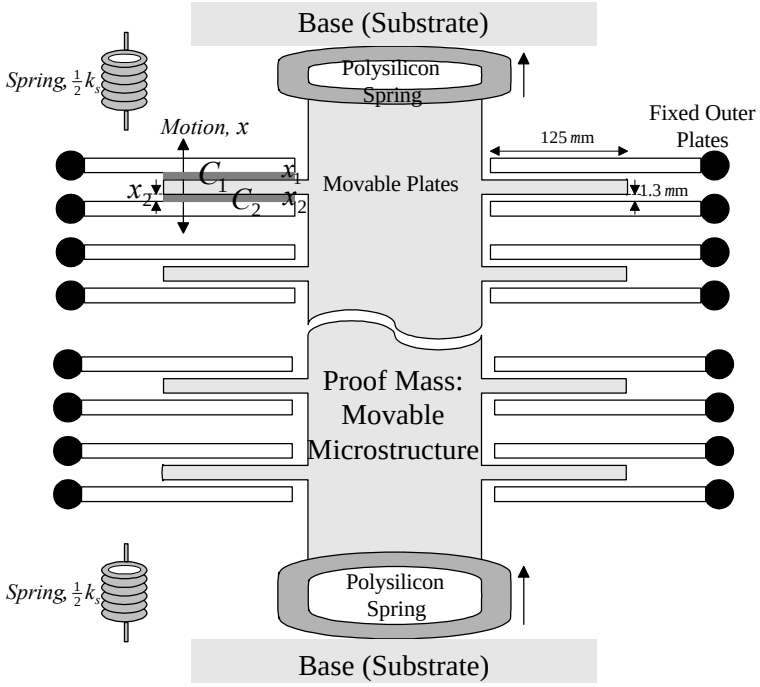


Figure 3.1.8. Accelerometer structure: proof mass, polysilicon springs, and sensing elements (fixed outer plates and central movable plates attached to the proof mass)

The proof mass (1.3 *m*m , 2 *m*m thick) has movable plates which are shown in Figure 3.1.8. The air capacitances C_1 and C_2 (capacitances between the movable plate and two stationary outer plates) are functions of the corresponding displacements x_1 and x_2 .

The parallel-plate capacitance is proportional to the overlapping area between the plates (125 *m*m 2 *m*m) and the displacement (up to 1.3 *m*m). In particular, neglecting the fringing effects (nonuniform distribution near the edges), the parallel-plate capacitance is

$$C = e\frac{A}{d} = e_A\frac{1}{d},$$

where e is the permittivity; A is the overlapping area; d is the displacement between plates; $e_A = eA$

If the acceleration is zero, the capacitances C_1 and C_2 are equal because $x_1 = x_2$ (in ADXL202 accelerometer, $x_1 = x_2 = 1.3\,m$m).

Thus, one has

$$C_1 = C_2,$$

where $C_1 = e_A \frac{1}{x_1}$ and $C_2 = e_A \frac{1}{x_2}$.

The proof mass (movable microstructure) displacement x results due to acceleration. If $x \quad 0$, we have the following expressions for capacitances

$$C_1 = e_A \frac{1}{x_1 + x} \text{ and } C_2 = e_A \frac{1}{x_2 \quad x} = e_A \frac{1}{x_1 \quad x}.$$

The capacitance difference is found to be

$$C = C_1 \quad C_2 = 2e_A \frac{x}{x^2 \quad x_1^2}.$$

Measuring C, one finds the displacement x by solving the following nonlinear algebraic equation

$$Cx^2 \quad 2e_A x \quad Cx_1^2 = 0.$$

For small displacements, neglecting the term Cx^2, one has

$$x \quad \frac{x_1^2}{2e_A} \quad C.$$

Hence, the displacement is proportional to the capacitance difference C.

For an ideal spring, Hook's law states that the spring exhibits a restoring force F_s which is proportional to the displacement x. Hence, we have the following formula

$F_s = k_s x$,

where k_s is the spring constant.

From Newton's second law of motion, neglecting friction, one writes

$$ma = m\frac{d^2x}{dt^2} = k_s x.$$

Thus, the displacement due to the acceleration is

$$x = \frac{m}{k_s} a,$$

while the acceleration, as a function of the displacement, is given as

$$a = \frac{k_s}{m} x.$$

Then, making use of the measured (calculated) C, the acceleration is found to be

$$a = \quad \frac{k_s x_1^2}{2me_A} \quad C.$$

Making use of Newton's second law of motion, we have

$$ma = m\frac{d^2x}{dt^2} = \underset{\text{spring force}}{f_s(x)} \ ,$$

where $f_s(x)$ is the spring restoring force which is a nonlinear function of the displacement, and $f_s(x) = k_{s1}x + k_{s2}x^2 + k_{s3}x^3$; k_{s1}, k_{s2} and k_{s3} are the spring constants.

Therefore, the following nonlinear equation results

$$ma = k_{s1}x + k_{s2}x^2 + k_{s3}x^3 .$$

Thus,

$$a = \frac{1}{m}\left(k_{s1}x + k_{s2}x^2 + k_{s3}x^3\right),$$

where x $\frac{x_1^2}{2e_A}$ C.

This equation can be used to calculate the acceleration a using the capacitance difference C.

Two beams (proof masses which are motion microstructures) can be placed orthogonally to measure the accelerations in the X and Y axis (ADXL250), as well as the movable plates can be mounted along the sides of the square beam (ADXL202). Figures 3.1.9 and 3.1.10 document the ADXL202 and ADXL250 accelerometers.

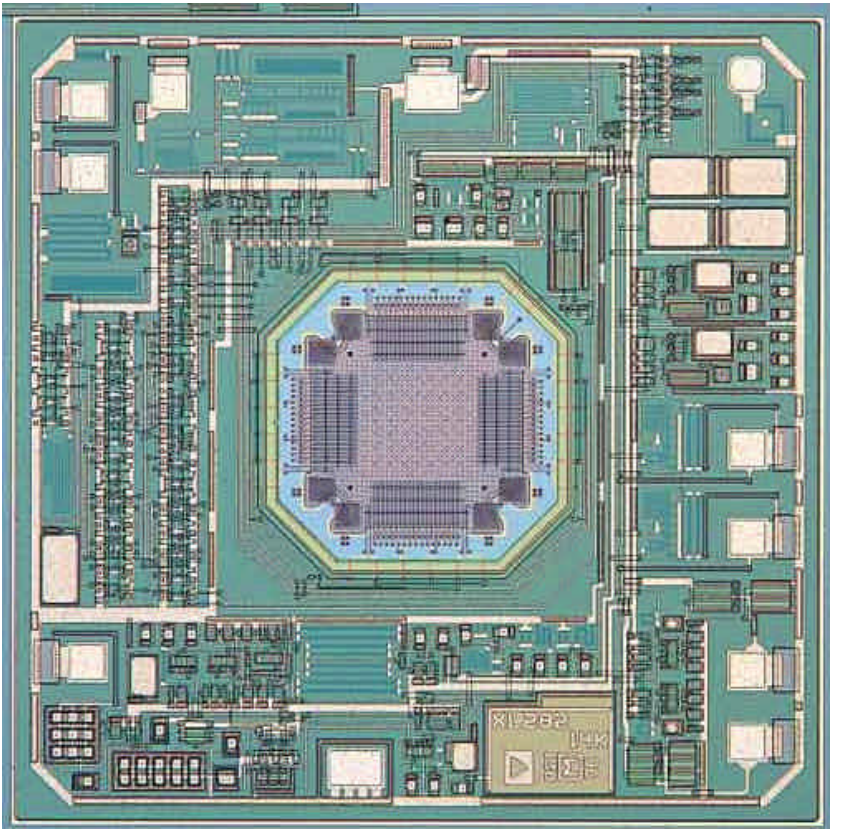


Figure 3.1.9. ADXL202 accelerometer: proof mass with fingers and ICs (courtesy of Analog Devices)

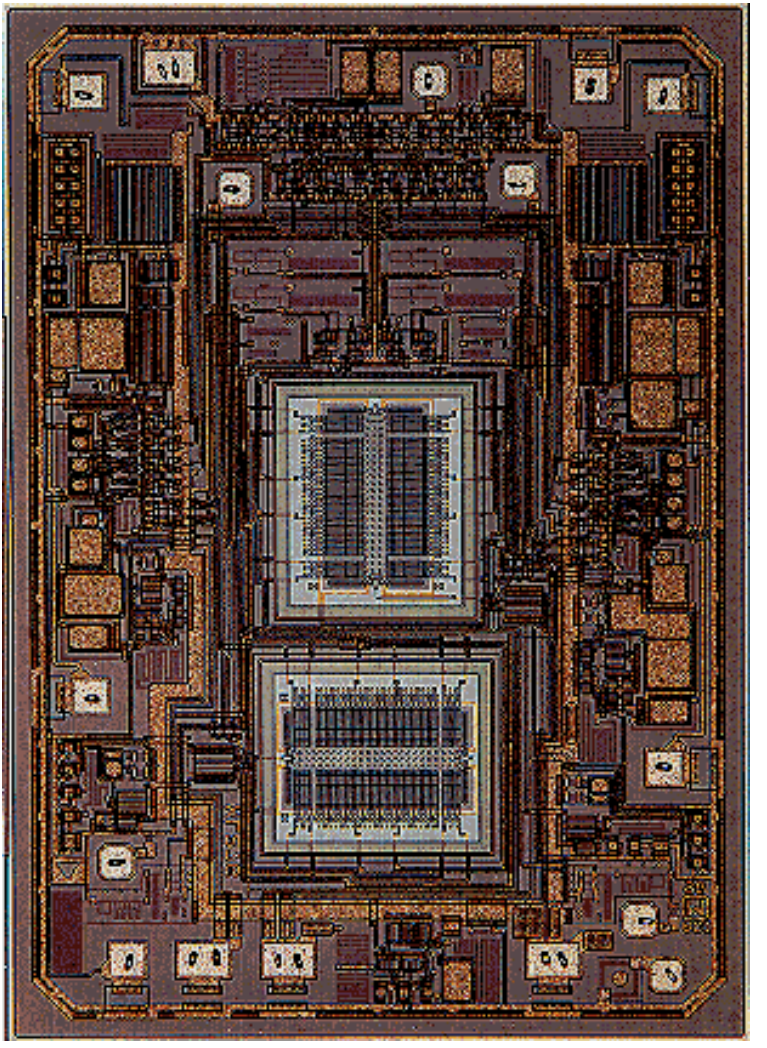


Figure 3.1.10.ADXL250 accelerometer: proof masses with fingers and ICs (courtesy of Analog Devices)

Responding to acceleration, the proof mass moves due to the mass of the movable microstructure (m) along X and Y axes relative to the stationary member (accelerometer). The motion of the proof mass is constrained, and the polysilicon springs hold the movable microstructure (beam). Assuming that the polysilicon springs and the proof mass obey Hook's and Newton's laws, it was shown that the acceleration is found using the following formula

$$a = \frac{k_s}{m} x .$$

The fixed outer plates are excited by two square wave 1 MHz signals of equal magnitude that are 180 degrees out of phase from each other. When the movable plates are centered between the fixed outer plates we have $x_1 = x_2$. Thus, the capacitance difference C and the output signal is zero. If the proof mass (movable microstructure) is displaced due to the acceleration, we have $C \neq 0$. Thus, the capacitance imbalance, and the amplitude of the output voltage is a function (proportional) to the displacement of the proof mass *x*. Phase demodulation is used to determine the sign (positive or negative) of acceleration. The ac signal is amplified by buffer amplifier and demodulated by a synchronous synchronized demodulator. The output of the demodulator drives the high-resolution duty cycle modulator. In particular, the filtered signal is converted to a PWM signal by the 14-bit duty cycle modulator. The zero acceleration produces 50% duty cycle. The PWM output fundamental period can be set from 0.5 to 10ms.

There is a wide range of industrial systems where smart integrated sensors are used. For example, accelerometers can be used for

1. active vibration control and diagnostics,
2. health and structural integrity monitoring,
3. internal navigation systems,
4. earthquake-actuated safety systems,
5. seismic instrumentation: monitoring and detection,
6. etc.

Current research activities in analysis, design, and optimization of flexible structures (aircraft, missiles, manipulators and robots, spacecraft, surface and underwater vehicles) are driven by requirements and standards which must be guaranteed. The vibration, structural integrity, and structural behavior are addressed and studied. For example, fundamental, applied, and experimental research in aeroelasticity and structural dynamics are conducted to obtain fundamental understanding of the basic phenomena involved in flutter, force and control responses, vibration, and control. Through optimization of aeroelastic characteristics as well as applying passive and active vibration control, the designer minimizes vibration and noise, and current research integrates development of aeroelastic models and diagnostics to predict stalled/whirl flutter, force and control responses, unsteady flight, aerodynamic flow, etc. Vibration control is a very challenging problem because the designer must account complex interactive physical phenomena (elastic theory, structural and continuum mechanics, radiation and transduction, wave propagation, chaos, et cetera). Thus, it is necessary to accurately measure the vibration, and the accelerometers, which allow one to measure the acceleration in the micro-g range, are used. The application of the MEMS-based accelerometers ensures small size, low cost,

ruggedness, hermeticity, reliability, and flexible interfacing with microcontrollers, microprocessors, and DSPs.

High-accuracy low-noise accelerometers can be used to measure the velocity and position. This provides the back-up in the case of the GPS system failures or in the dead reckoning applications (the initial coordinates and speed are assumed to be known). Measuring the acceleration, the velocity and position in the *xy* plane are found using integration. In particular,

$$v_x(t) = \int_{t_0}^{t_f} a_x(t)dt\,,\ v_y(t) = \int_{t_0}^{t_f} a_y(t)dt\,,$$

$$x_x(t) = \int_{t_0}^{t_f} v_x(t)dt\,,\ x_y(t) = \int_{t_0}^{t_f} v_y(t)dt\,.$$

The Analog Devices data for *i*MEMS accelerometers ADXL202/ADXL210 and ADXL150/ADXL250 are given below (courtesy of Analog Devices).

It is important to emphasize that microgyroscope have been designed, fabricated, and deployed using the similar technology as *i*MEMS accelerometers. In particular, using the difference capacitance (between the movable rotor and stationary stator plates), the angular acceleration is measured. The butterfly-shaped polysilicon rotor suspended above the substrate, and Figure 3.1.11 illustrates the microgyroscope.

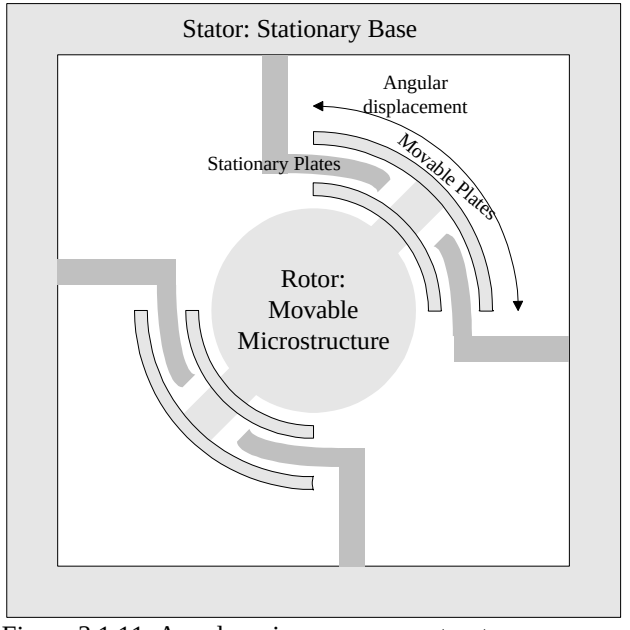


Figure 3.1.11. Angular microgyroscope structure

Using the experimental data (input-output dynamic behavior and Bode plots), the mathematical model of microaccelerometers is obtained in the form of ordinary differential equations, and the coefficients (accelerometer parameters) are identified. The dominant microaccelerometer dynamics is described by a system of six linear differential equations

$$\frac{dx}{dt} = Ax + Bu,\ y = Cx,$$

where the matrices of coefficients are

$$A = \begin{matrix} 2.6\ 10^{4} & 2.7\ 10^{10} & 4.2\ 10^{14} & 1.5\ 10^{20} & 9\ 10^{23} & 3.7\ 10^{27} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{matrix}, \quad B = \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix},$$

$$C = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 3.7\ 10^{27}].$$

The accelerometer output, which is the measured acceleration a, was denoted as y, $y = a$. It is evident that the acceleration is a function of the state variable x_6. All other five states model the proof mass (motion microstructure) and microICs (oscillator, demodulator, modulator, filter, et cetera) dynamics.

The eigenvalues are found to be

$$5.9\ \ 10^{3} \pm i1.4\ \ 10^{5}, \quad 4.2\ \ 10^{3} \pm i8.8\ \ 10^{4}, \quad 3\ \ 10^{3} \pm i4\ \ 10^{3}.$$

This mathematical model of the microaccelerometer can be used in systems analysis, diagnostics, and design of a wide variety of systems where *i*MEMS are used.

Low Cost ±2 *g*/±10 *g* Dual Axis *i*MEMS® Accelerometers with Digital Output

ADXL202/ADXL210

FEATURES

2-Axis Acceleration Sensor on a Single IC Chip
Measures Static Acceleration as Well as Dynamic Acceleration
Duty Cycle Output with User Adjustable Period
Low Power <0.6 mA
Faster Response than Electrolytic, Mercury or Thermal Tilt Sensors
Bandwidth Adjustment with a Single Capacitor Per Axis
5 m*g* Resolution at 60 Hz Bandwidth
+3 V to +5.25 V Single Supply Operation
1000 *g* Shock Survival

APPLICATIONS

2-Axis Tilt Sensing
Computer Peripherals
Inertial Navigation
Seismic Monitoring
Vehicle Security Systems
Battery Powered Motion Sensing

GENERAL DESCRIPTION

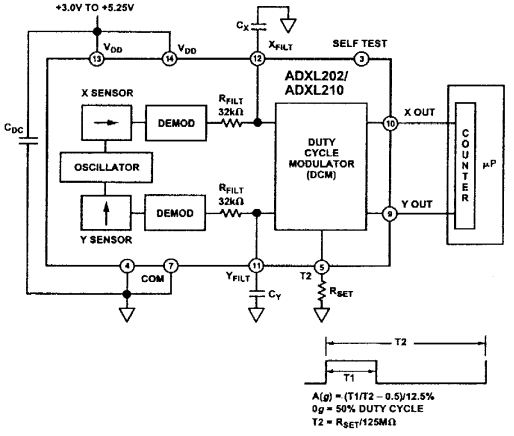
The ADXL202/ADXL210 are low cost, low power, complete 2-axis accelerometers with a measurement range of either ±2 *g*/±10 *g*. The ADXL202/ADXL210 can measure both dynamic acceleration (e.g., vibration) and static acceleration (e.g., gravity).

The outputs are digital signals whose duty cycles (ratio of pulsewidth to period) are proportional to the acceleration in each of the 2 sensitive axes. These outputs may be measured directly with a microprocessor counter, requiring no A/D converter or glue logic. The output period is adjustable from 0.5 ms to 10 ms via a single resistor (R_{SET}). If a voltage output is desired, a voltage output proportional to acceleration is available from the X_{FILT} and Y_{FILT} pins, or may be reconstructed by filtering the duty cycle outputs.

The bandwidth of the ADXL202/ADXL210 may be set from 0.01 Hz to 5 kHz via capacitors C_X and C_Y. The typical noise floor is 500 µ*g*/$\sqrt{Hz}$ allowing signals below 5 m*g* to be resolved for bandwidths below 60 Hz.

The ADXL202/ADXL210 is available in a hermetic 14-lead Surface Mount CERPAK, specified over the 0°C to +70°C commercial or –40°C to +85°C industrial temperature range.

FUNCTIONAL BLOCK DIAGRAM



*i*MEMS is a registered trademark of Analog Devices, Inc.

REV. B

Information furnishisd by Analog Devices is believed to be accurate and reliable. However, no responsibility is assumed by Analog Devices for its use, nor for any infringements of patents or other rights of third parties which may result from its use. No license is granted by implication or otherwise under any patent or patent rights of Analog Devices.

One Technology Way, P.O. Box 9106, Norwood, MA 02062-9106, U.S.A.
Tel: 781/329-4700 World Wide Web Site: http://www.analog.com
Fax: 781/326-8703
© Analog Devices, Inc., 1999

ADXL202/ADXL210–SPECIFICATIONS

($T_A = T_{MIN}$ to T_{MAX}, T_A = +25°C for J Grade only, V_{DD} = +5 V, R_{SET} = 125 kΩ, Acceleration = 0 *g*, unless otherwise noted)

		ADXL202/JQC/AQC			ADXL210/JQC/AQC			
Parameter	**Conditions**	**Min**	**Typ**	**Max**	**Min**	**Typ**	**Max**	**Units**
SENSOR INPUT	Each Axis							
Measurement Range[1]		±1.5	±2		±8	±10		*g*
Nonlinearity	Best Fit Straight Line		0.2			0.2		% of FS
Alignment Error[2]			±1			±1		Degrees
Alignment Error	X Sensor to Y Sensor		±0.01			±0.01		Degrees
Transverse Sensitivity[3]			±2			±2		%
SENSITIVITY	Each Axis							
Duty Cycle per *g*	T1/T2 @ +25°C	10	12.5	15	3.2	4.0	4.8	%/*g*
Sensitivity, Analog Output	At Pins X_{FILT}, Y_{FILT}		312			100		mV/*g*
Temperature Drift[4]	Δ from +25°C		±0.5			±0.5		% Rdg
ZERO *g* BIAS LEVEL	Each Axis							
0 *g* Duty Cycle	T1/T2	25	50	75	42	50	58	%
Initial Offset			±2			±2		*g*
0 *g* Duty Cycle vs. Supply			1.0	4.0		1.0	4.0	%/V
0 *g* Offset vs. Temperature[4]	Δ from +25°C		2.0			2.0		mg/°C
NOISE PERFORMANCE								
Noise Density[5]	@ +25°C		500	1000		500	1000	$\mu g/\sqrt{Hz}$
FREQUENCY RESPONSE								
3 dB Bandwidth	Duty Cycle Output		500			500		Hz
3 dB Bandwidth	At Pins X_{FILT}, Y_{FILT}		5			5		kHz
Sensor Resonant Frequency			10			14		kHz
FILTER								
R_{FILT} Tolerance	32 kΩ Nominal		±15			±15		%
Minimum Capacitance	At X_{FILT}, Y_{FILT}	1000			1000			pF
SELF TEST								
Duty Cycle Change	Self-Test "0" to "1"		10			10		%
DUTY CYCLE OUTPUT STAGE								
F_{SET}			125 MΩ/R_{SET}			125 MΩ/R_{SET}		
F_{SET} Tolerance	R_{SET} = 125 kΩ	0.7		1.3	0.7		1.3	kHz
Output High Voltage	I = 25 µA	V_S – 200 mV			V_S – 200 mV			mV
Output Low Voltage	I = 25 µA			200			200	mV
T2 Drift vs. Temperature			35			35		ppm/°C
Rise/Fall Time			200			200		ns
POWER SUPPLY								
Operating Voltage Range		3.0		5.25	2.7		5.25	V
Specified Performance		4.75		5.25	4.75		5.25	V
Quiescent Supply Current			0.6	1.0		0.6	1.0	mA
Turn-On Time[6]	To 99%		160 C_{FILT} + 0.3			160 C_{FILT} + 0.3		ms
TEMPERATURE RANGE								
Operating Range	JQC	0		+70	0		+70	°C
Specified Performance	AQC	–40		+85	–40		+85	°C

NOTES

[1]For all combination of offset no sensitivity variation.

[2]Alignment error is specified as the angle between the true and indicated axis of sensitivity.

[3]Transverse sensitivity is the algebraic non of the alignment and the inherent sensitivity errors.

[4]Specification refers to the maximum change in parameter from its initial at +25°C to its worst case value at T_{MIN} T_{MAX}.

[5]Noose density ($\mu g/\sqrt{Hz}$) is the average noise at any frequency in the bandwith of the part.

[6]C_{FILT} in µF. Addition of filter capacitor will increase turn on time. Please see the Application section on power cycling.

All min and max specifications are guaranteed. Typical specifications are not tested or guaranteed.

Specifications subject to change without notice.

ABSOLUTE MAXIMUM RATINGS*

Acceleration (Any Axis, Unpowered for 0.5 ms) . . . . . 1000 *g*
Acceleration (Any Axis, Powered for 0.5 ms) . . . . . 500 *g*
$+V_S$ . . . . . –0.3 V to +7.0 V
Output Short Circuit Duration
(Any Pin to Common) . . . . . Indefinite
Operating Temperature . . . . . –55°C to +125°C
Storage Temperature . . . . . –65°C to +I 50°C

*Stresses above those listed under Absolute Maximum Ratings may cause permanent damage to the device. This is a stress rating only; the functional operation of the device at these or any other conditions above those indicated in the operational sections of this specification is not implied. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

Drops onto hard surfaces can cause shocks of greater than 1000 *g* and exceed the absolute maximum rating of the device. Care should be exercised in handling to avoid damage.

PIN FUNCTION DESCRIPTIONS

Pin	Name	Description
1	NC	Not Connect
2	V_{TP}	Test Point, Do Not Connect
3	ST	Self Test
4	COM	Common
5	T2	Connect R_{SET} to Set T2 Period
6	NC	No Connect
7	COM	Common
8	NC	No Connect
9	Y_{OUT}	Y Axis Duty Cycle Output
10	X_{OUT}	X Axis Duty Cycle Output
11	Y_{FILT}	Connect Capacitor for Y Filter
12	X_{FILT}	Connect Capacitor for X Filter
13	V_{DD}	+3 V to +5.25 V, Connect to 14
14	V_{DD}	+3 V to +5.25 V, Connect to 13

PACKAGE CHARACTERISTICS

Package	θ_{JA}	θ_{JC}	Device Weight
14-Lead CERPAK	110°C/W	30°C/W	5 Grams

PIN CONFIGURATION

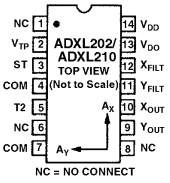


Figure 1 shows the response of the ADXL202 to the Earth's gravitational field. The output values shown are nominal. They are presented to show the user what type of response to expect from each of the output pins due to changes in orientation with respect to the Earth. The ADXL210 reacts similarly with output changes appropriate to its scale.

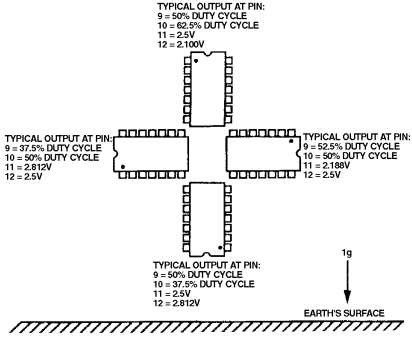


Figure 1. ADXL202/ADXL210 Nominal Response Due to Gravity

ORDERING GUIDE

Model	*g* Range	Temperature Range	Package Description	Package Option
ADXL202JQC	±2	0°C to +70°C	14-Lead CERPAK	QC-14
ADXL202AQC	±2	–40°C to +85°C	14-Lead CERPAK	QC-14
ADXL210JQC	±10	0°C to +70°C	14-Lead CERPAK	QC-14
ADXL210AQC	±10	–40°C to +85°C	14-Lead CERPAK	QC-14

CAUTION

ESD (electrostatic discharge) sensitive device. Electrostatic charges as high as 4000 V readily accumulate on the human body and test equipment and can discharge without detection. Although the ADXL202/ADXL210 features proprietary ESD protection circuitry, permanent damage may occur on devices subjected to high energy electrostatic discharges. Therefore, proper ESD precautions are recommended to avoid performance degradation or loss of functionality.

TYPICAL CHARACTERISTICS (@ +25°C R_{SET} = 125 kΩ, V_{DD} = +5 V, unless otherwise noted)

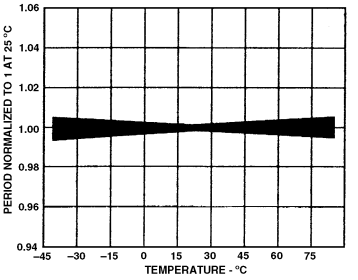


Figure 2. Normalized DCM Period (T2) vs. Temperature

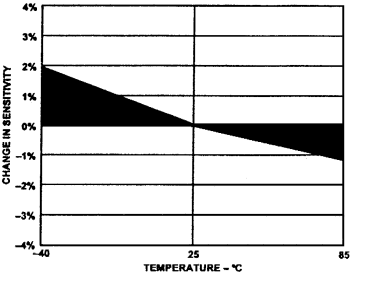


Figure 5. Typical X Axis Sensitivity Drift Due to Temperature

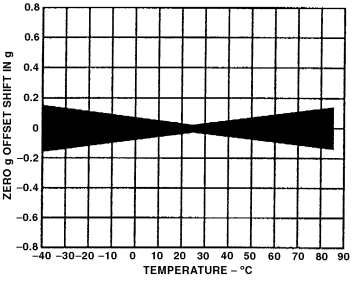


Figure 3. Typical Zero g Offset vs. Temperature

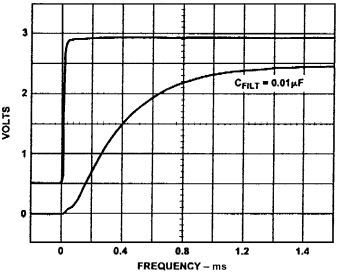


Figure 6. Typical Turn-On Time

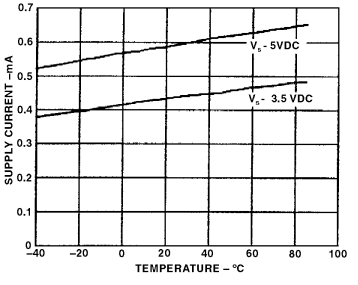


Figure 4. Typical Supply Current vs. Temperature

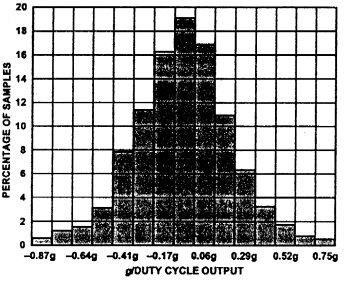


Figure 7. Typical Zero g Distribution at +25°C

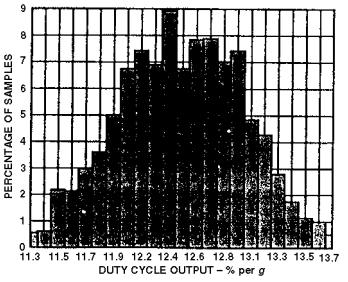


Figure 8. Typical Sensitivity per g at +25°C

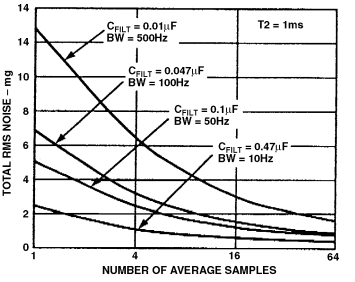


Figure 10. Typical Noise at Digital Outputs

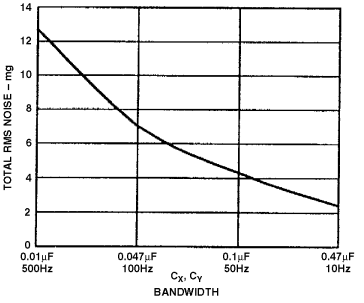


Figure 9. Typical Noise at X_{FILT} Output

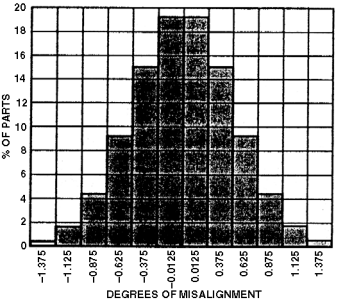


Figure 11. Rotational Die Alignment

DEFINITIONS

T1	Length of the "on" portion of the cycle.
T2	Length of the total cycle.
Duty Cycle	Ratio of the "on" time (T1) of the cycle to the total cycle (T2). Defined as TIM for the ADXL202/ADXL210.
Pulsewidth	Time period of the "on" pulse. Defined as T1 for the ADXL202/ADXL210.

THEORY OF OPERATION

The ADXL202/ADXL210 are complete dual axis acceleration measurement systems on a single monolithic IC. They contain a polysilicon surface-micromachined sensor and signal conditioning circuitry to implement an open loop acceleration measurement architecture. For each axis, an output circuit converts the analog signal to a duty cycle modulated (DCM) digital signal that can be decoded with a counter/timer port on a microprocessor cessor. The ADXL202/ADXL210 are capable of measuring both positive and negative accelerations to a maximum level of ± 2 *g* or ± 10 *g*. The accelerometer measures static acceleration forces such as gravity, allowing it to be used as a tilt sensor.

The sensor is a surface micromachined polysilicon structure built on top of the silicon wafer. Polysilicon springs suspend the structure over the surface of the wafer and provide a resistance against acceleration forces. Deflection of the structure is measured using a differential capacitor that consists of independent fixed plates and central plates attached to the moving mass. The fixed plates are driven by 180° out of phase square waves. An acceleration will deflect the beam and unbalance the differential capacitor, resulting in an output square wave whose amplitude is proportional to acceleration. Phase sensitive demodulation techniques are then used to rectify the signal and determine the direction of the acceleration.

The output of the demodulator drives a duty cycle modulator (DCM) stage through a 32 kΩ resistor. At this point a pin is available on each channel to allow the user to set the signal bandwidth of the device by adding a capacitor. This filtering improves measurement resolution and helps prevent aliasing.

After being low-pass filtered, the analog signal is converted to a duty cycle modulated signal by the DCM stage. A single resistor sets the period for a complete cycle (T2), which can be set between 0.5 ms and 10 ms (see Figure 12). A 0 *g* acceleration produces a nominally 50% duty cycle. The acceleration signal can be determined by measuring the length of the T1 and T2 pulses with a counter/timer or with a polling loop using a low cost microcontroller

An analog output voltage can be obtained either by buffering the signal from the X_{FILT} and Y_{FILT} pin, or by passing the duty cycle signal through an RC filter to reconstruct the dc value.

The ADXL202/ADXL210 will operate with supply voltages as low as 3.0 V or as high as 5.25 V.

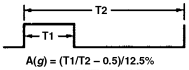


Figure 12. Typical Output Duty Cycle

APPLICATIONS

POWER SUPPLY DECOUPLING

For most applications a single 0. 1 µF capacitor, C_{DC}, will adequately decouple the accelerometer from signal and noise on the power supply. However, in some cases, especially where digital devices such as microcontrollers share the same power supply, digital noise on the supply may cause interference on the ADXL202/ ADXL210 output. This is often observed as a slowly undulating fluctuation of voltage at X_{FILT} and Y_{FILT}. If additional decoupling is needed, a 100 Ω (or smaller) resistor or ferrite beads, may be inserted in the ADXL202/ADXL210's supply line.

DESIGN PROCEDURE FOR THE ADXL202/ADXL210

The design procedure for using the ADXL202/ADXL210 with a duty cycle output involves selecting a duty cycle period and a filter capacitor. A proper design will take into account the Application requirements for bandwidth, signal resolution and acquisition time, as discussed in the following sections.

V_{DD}

The ADXL202/ADXL210 have two power supply (V_{DD}) Pins: 13 and 14. These two pins should be connected directly together.

COM

The ADXL202/ADXL210 have two commons, Pins 4 and 7. These two pins should be connected directly together and Pin 7 grounded.

V_{TP}

This pin is to be left open; make no connections of any kind to this pin.

Decoupling Capacitor C_{DC}

A 0.1 µF Capacitor is recommended from Von to COM for power supply decoupling.

ST

The ST pin controls the self-test feature. When this pin is set to V_{DD}, an electrostatic force is exerted on the beam of the accelerometer. The resulting movement of the beam allows the user to test if the accelerometer is functional. The typical change in output will be 10% at the duty cycle outputs (corresponding to 800 mg). This pin may be left open circuit or connected to common in normal use.

Duty Cycle Decoding

The ADXL202/ADXL210's digital output is a duty cycle modulator. Acceleration is proportional to the ratio T1/T2. The nominal output of the ADXL202 is:

$$0\ g = 50\%\ Duty\ Cycle$$

Scale factor is 12.5% Duty Cycle Change per *g*

The nominal output of the ADXL210 is:

$$0\ g = 50\%\ Duty\ Cycle$$

Scale factor is 4% Duty Cycle Change per *g*

These nominal values are affectcd by the initial tolerance of the device including zero *g* offset error and sensitivity error.

T2 does not have to be measured for every measurement cycle. It need only be updated to account for changes due to temperature, (a relatively slow process). Since the T2 time period is shared by both X and Y channels, it is necessary only to measure it on one channel of the ADXL202/ADXL210. Decoding algorithms for various microcontrollers have been developed. Consult the appropriate Application Note.

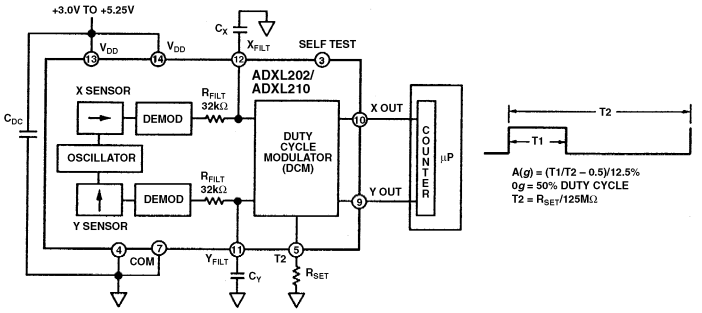


Figure 13. Block Diagram

Setting the Bandwidth Using C_X and C_Y

The ADXL202/ADXL210 have provisions for bandlimiting the X_{FILT} and Y_{FILT} pins. Capacitors must be added at these pins to implement low-pass filtering for antialiasing and noise reduction. The equation for the 3 dB bandwidth is:

$$F_{-3\,dB} = \frac{1}{\left(2\pi(32\,k\Omega) \times C(x, y)\right)}$$

or, more simply, $F_{-3\,dB} = \frac{5\mu F}{C_{(X,Y)}}$

The tolerance of the internal resistor (R_{FILT}) can vary as much as ±25% of its nominal value of 32 kΩ; so the bandwidth will vary accordingly. A minimum capacitance of 1000 pF for $C_{(X,Y)}$ is required in all cases.

Table I. Filter Capacitor Selection, C_X and C_Y

Bandwidth	Capacitor Value
10 Hz	0.47 µF
50 Hz	0.10 µF
100 Hz	0.05 µF
200 Hz	0.027 µF
500 Hz	0.01 µF
5 kHz	0.001 µF

Setting the DCM Period with R_{SET}

The period of the DCM output is set for both channels by a single resistor from R_{SET} to ground. The equation for the period is:

$$T2 = \frac{R_{SET}\,(\Omega)}{125\ M\Omega}$$

A 125 kΩ resistor will set the duty cycle repetition rate to approximately 1 kHz, or 1 ms. The device is designed to operate at duty cycle periods between 0.5 ins and 10 ms.

Table II. Resistor Values to Set T2

T2	R_{SET}
1 ms	125 kΩ
2 ins	250 kΩ
5 ms	625 kΩ
10 ms	1.25 MΩ

Note that the R_{SET} should always be included, even if only an analog output is desired. Use an R_{SET} value between 500 kΩ and 2 MΩ when taking the output from X_{FILT} or Y_{FILT}. The R_{SET} resistor should be place close to the T2 Pin to minimize parasitic capacitance at this node.

Selecting the Right accelerometer

For most tilt sensing applications the ADXL202 is the most appropriate accelerometer. Its higher sensitivity (12.5%/*g* allows the user to use a lower speed counter for PWM decoding while maintaining high resolution. The ADXL210 should be used in applications where accelerations of greater than ±2 *g* are expected.

MICROCOMPUTER INTERFACES

The ADXL202/ADXL210 were specifically designed to work with low cost microcontrollers. Specific code sets, reference designs, and application notes are available from the factory. This section will outline a general design procedure and discuss the various trade-offs that need to be considered.

The designer should have some idea of the required performance of the system in terms of:

Resolution: the smallest signal change that needs to be detected.

Bandwidth: the highest frequency that needs to be detected.

Acquisition Time: the time that will be available to acquire the signal on each axis.

These requirements will help to determine the accelerometer bandwidth, the speed of the microcontroller clock and the length of the T2 period.

When selecting a microcontroller it is helpful to have a counter timer port available. The microcontroller should have provisions for software calibration. While the ADXL202/ADXL210 are highly accurate accelerometers, they have a wide tolerance for

initial offset. The easiest way to null this offset is with a calibration factor saved on the microcontroller or by a user calibration for zero *g*. In the case where the offset is calibrated during manufacture, there are several options, including external EEPROM and microcontrollers with "one-time programmable" features.

DESIGN TRADE-OFFS FOR SELECTING FILTER CHARACTERISTICS: THE NOISE/BW TRADE-OFF

The accelerometer bandwidth selected will determine the measurement resolution (smallest detectable acceleration). Filtering can be used to lower the noise floor and improve the resolution of the accelerometer. Resolution is dependent on both the analog filter bandwidth at X_{FILT} and Y_{FILT} and on the speed of the microcontroller counter.

The analog output of the ADXL202/ADXL210 has a typical bandwidth of 5 kHz, much higher than the duty cycle stage is capable of converting. The user must filter the signal at this point to limit aliasing errors. To minimize DCM errors the analog bandwidth should be less than 1/10 the DCM frequency. Analog bandwidth may be increased to up to 1/2 the DCM frequency in many applications. This will result in greater dynamic error generated at the DCM.

The analog bandwidth may be further decreased to reduce noise and improve resolution. The ADXL202/ADXL210 noise has the characteristics of white Gaussian noise that contributes equally at all frequencies and is described in terms of µg per root Hz; i.e., the noise is proportional to the square root of the handwidth of the accelerometer. It is recommended that the user limit bandwidth to the lowest frequency needed by the application, to maximize the resolution and dynamic range of the accelerometer.

With the single pole roll-off characteristic, the typical noise of the ADXL202/ADXL210 is determined by the following equation:

$$Noise(rms) = \left(500\ \mu g/\sqrt{Hz}\right) \times \left(\sqrt{BW \times 1.5}\right)$$

At 100 Hz the noise will be:

$$Noise(rms) = \left(500 \mu g/\sqrt{Hz}\right) \times \left(\sqrt{100 \times (1.5)}\right) = 6.12\ mg$$

Often the peak value of the noise is desired. Peak-to-peak noise can only be estimated by statistical methods. Table III is useful for estimating the probabilities of exceeding various peak values, given the rms value.

Table III. Estimation of Peak-to-Peak Noise

Nominal Peak-to-Peak Value	% of Time that Noise Will Exceed Nominal Peak-to-Peak Value
2.0 × rms	32%
4.0 × rms	4.6%
6.0 × rms	0.27%
8.0 × rms	0.006%

The peak-to-peak noise value will give the best estimate of the uncertainty in a single measurement.

Table IV gives typical noise output of the ADXL202/ADXL210 for various C_X and C_Y values.

Table IV. Filter Capacitor Selection, C_X and C_Y

Bandwidth	C_X, C_Y	rms Noise	Peak-to-Peak Noise Estimate 95% Probability (rms × 4)
10 Hz	0.47 µF	1.9 mg	7.6 mg
50 Hz	0.10 µF	4.3 mg	17.2 mg
100 Hz	0.05 µF	6.1 mg	24.4 mg
200 Hz	0.027 µF	8.7 mg	35.8 mg
500 Hz	0.01 µF	13.7 mg	54.8 mg

CHOOSING T2 AND COUNTER FREQUENCY: DESIGN TRADE-OFFS

The noise level is one determinant of accelerometer resolution. The second relates to the measurement resolution of the counter when decoding the duty cycle output.

The ADXL202/ADXL210's duty cycle converter has a resolution of approximately 14 bits; better resolution than the accelerometer itself. The actual resolution of the acceleration signal is, however, limited by the time resolution of the counting devices used to decode the duty cycle. The faster the counter clock, the higher the resolution of the duty cycle and the shorter the T2 period can be for a given resolution. The following table shows some of the trade-offs. It is important to note that this is the resolution due to the microprocessors's counter. It is probable that the accelerometer's noise floor may set the lower limit on the resolution as discussed in the previous section.

Table V. Trade-offs Between Microcontroller Counter Rate, T2 Period and Resolution of Duty Cycle Modulator

T2(ms)	R_{SET} (kΩ)	ADXL202/ ADXL210 Sample Rate	Counter-Clock Rate (MHz)	Counts per T2 Cycle	Counts per *g*	Resolution (mg)
1.0	124	1000	2.0	2000	250	4.0
1.0	124	1000	1.0	1000	125	8.0
1.0	124	1000	0.5	500	62.5	16.0
5.0	625	200	2.0	10000	1250	0.8
5.0	625	200	1.0	5000	625	1.6
5.0	625	200	0.5	2500	312.5	3.2
10.0	1250	100	2.0	20000	2500	0.4
10.0	1250	100	1.0	10000	1250	0.8
10.0	1250	100	0.5	5000	625	1.6

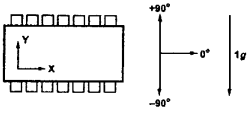
STRATEGIES FOR USING THE DUTY CYCLE OUTPUT WITH MICROCONTROLLERS

Application notes outlining various strategies for using the duty cycle output with low cost microcontrollers are available from the factory.

USING THE ADXL202/ADXL210 AS A DUAL AXIS TILT SENSOR

One of the most popular applications of the ADXL202/ADXL210 is tilt measurement. An accelerometer uses the force of gravity as an input vector to determine orientation of an object in space.

An accelerometer is most sensitive to tilt when its sensitive axis is perpendicular to the force of gravity, i.e., parallel to the earth's surface. At this orientation its sensitivity to changes in tilt is highest. When the accelerometer is oriented on axis to gravity, i. e., near its +1 *g* or −1 *g* reading, the change in output acceleration per degree of tilt is negligible. When the accelerometer is perpendicular to gravity, its output will change nearly 17.5 mg per degree of tilt, but at 45° degrees it is changing only at 12.2 mg per degree and resolution declines. The following table illustrates the changes in the X and Y axes as the device is tilted ±90° through gravity.



X AXIS ORIENTATION TO HORIZON (°)	X OUTPUT		Y OUTPUT (*g*)	
	X OUTPUT (*g*)	D PER DEGREE OF TILT (mg)	Y OUTPUT (*g*)	Δ PER DEGREE OF TILT (mg)
–90	–1.000	–0.2	0.000	17.5
–75	–0.966	4.4	0.259	16.9
–60	–0.866	8.6	0.500	15.2
–45	–0.707	12.2	0.707	12.4
–30	–0.500	15.0	0.866	8.9
–15	–0.259	16.8	0.966	4.7
0	0.000	17.5	1.000	0.2
15	0.259	16.9	0.966	–4.4
30	0.500	15.2	0.866	–8.6
45	0.707	12.4	0.707	–12.2
60	0.866	8.9	0.500	–15.0
75	0.966	4.7	0.259	–16.8
90	1.000	0.2	0.000	–17.5

Figure 14. How the X and Y Axes Respond to Changes in Tilt

A DUAL AXIS TILT SENSOR: CONVERTING ACCELERATION TO TILT

When the accelerometer is oriented so both its X and Y axes are parallel to the earth's surface it can be used as a two axis tilt sensor with a roll and a pitch axis. Once the output signal from the accelerometer has been converted to an acceleration that varies between −1 *g* and +1 *g*, the output tilt in degrees is calculated as follows:

$$Pitch = ASIN\ (Ax/1\ g)$$

$$Roll = ASIN\ (Ay/1\ g)$$

Be sure to account for overranges. It is possible for the accelerometers to output a signal greater than ± 1 *g* due to vibration, shock or other accelerations.

MEASURING 360° OF TILT

It is possible to measure a full 360° of orientation through gravity by using two accelerometers oriented perpendicular to one another (see Figure 15). When one sensor is reading a maximum change in output per degree, the other is at its minimum.

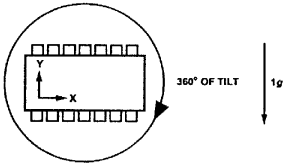


Figure 15. Using a Two-Axis Accelerometer to Measure 360° of Tilt

USING THE ANALOG OUTPUT

The ADXL202/ADXL210 was specifically designed for use with its digital outputs, but has provisions to provide analog outputs as well.

Duty Cycle Filtering

An analog output can be reconstructed by filtering the duty cycle output. This technique requires only passive components. The duty cycle period (T2) should be set to 1 ms. An RC filter with a 3 dB point at least a factor of 10 less than the duty cycle frequency is connected to the duty cycle output. The filter resistor should be no less than 100 kΩ to prevent loading of the output stage. The analog output signal will be ratiometric to the supply voltage. The advantage of this method is an output scale factor of approximately double the analog output. Its disadvantage is that the frequency response will be lower than when using the X_{FILT}, Y_{FILT} output.

X_{FILT}, Y_{FILT} Output

The second method is to use the analog output present at the X_{FILT} and Y_{FILT} pin. Unfortunately, these pins have a 32 kΩ output impedance and are not designed to drive a load directly. An op amp follower may be required to buffer this pin. The advantage of this method is that the full 5 kHz bandwidth of the accelerometer is available to the user. A capacitor still must be added at this point for filtering. The duty cycle converter should be kept running by using R_{SET} <10 MΩ. Note that the accelerometer offset and sensitivity are ratiometric to the supply voltage. The offset and sensitivity are nominally:

0 *g* Offset = $V_{DD}/2$	2.5 V at +5 V
ADXL202 Sensitivity = (60 mV × V_S)/g	300 mV/*g* at +5 V, V_{DD}
ADXL2l0 Sensitivity = (20 mV × V_S)/g	100 mV/*g* at +5 V, V_{DD}

USING THE ADXL202/ADXL210 IN VERY LOW POWER APPLICATIONS

An application note outlining low power strategies for the ADXL202/ADXL210 is available. Some key points are presented here. It is possible to reduce the ADXL202/ADXL210's average current from 0.6 mA to less than 20 µA by using the following techniques:

1. Power Cycle the accelerometer.
2. Run the accelerometer at a Lower Voltage, (Down to 3 V).

Power Cycling with an External A/D

Depending on the value of the X_{FILT} capacitor, the ADXL202/ADXL210 is capable of turning on and giving a good reading in 1.6 ms. Most microcontroller based A/Ds can acquire a reading in another 25 µs. Thus it is possible to turn on the ADXL202/ADXL210 and take a reading in <2 ms. If we assume that a 20 Hz sample rate is sufficient, the total current required to take 20 samples is 2 ms × 20 samples/s × 0.6 mA = 24 µA average current. Running the part at 3 V will reduce the supply current from 0.6 mA to 0.4 mA, bringing the average current down to 16 µA.

The A/D should read the analog output of the ADXL202/ADXL210 at the X_{FILT} and Y_{FILT} pins. A buffer amplifier is recommended, and may be required in any case to amplify the analog Output to give enough resolution with an 8-bit to 10-bit converter.

Power Cycling When Using the Digital Output

An alternative is to run the microcontroller at a higher clock rate and put it into shutdown between readings, allowing the use of the digital output. In this approach the ADXL202/ADXL210 should be set at its fastest sample rate (T2 = 0.5 ms), with a 500 Hz filter at X_{FILT} and Y_{FILT}. The concept is to acquire a reading as quickly as possible and then shut down the ADXL202/ADXL210 and the microcontroller until the next sample is needed.

In either of the above approaches, the ADXL202/ADXL210 can be turned on and off directly using a digital port pin on the microcontroller to power the accelerometer without additional components. The port should be used to switch the common pin of the accelerometer so the port pin is "pulling down."

CALIBRATING THE ADXL202/ADXL210

The initial value of the offset and scale factor for the ADXL202/ADXL210 will require calibration for applications such as tilt measurement. The ADXL202/ADXL210 architecture has been designed so that these calibrations take place in the software of the microcontroller used to decode the duty cycle signal. Calibration factors can be stored in EEPROM or determined at turn-on and saved in dynamic memory.

For low *g* applications, the force of gravity is the most stable, accurate and convenient acceleration reference available. A reading of the 0 *g* point can be determined by orientating the device parallel to the earth's surface and then reading the output.

A more accurate calibration method is to make a measurements at +1 *g* and −1 *g*. The sensitivity can be determined by the two measurements.

To calibrate, the accelerometer's measurement axis is pointed directly at the earth. The 1 *g* reading is saved and the sensor is turned 180° to measure −1 *g*. Using the two readings, the sensitivity is:

Let A = Accelerometer output with axis oriented to +1 *g*

Let B = Accelerometer output with axis oriented to −1 *g then*:

$$Sensitivity = [A - B]/2\ g$$

For example, if the +1 g reading (A) is 55% duty cycle and the −1 *g* reading (B) is 32% duty cycle, then:

$$Sensitivity = [55\% - 32\%]/2\ g = 11.5\%/g$$

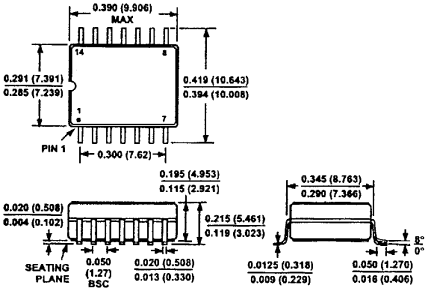
These equations apply whether the output is analog, or duty cycle.

Application notes outlining algorithms for calculating acceleration from duty cycle and automated calibration routines are available from the factory.

OUTLINE DIMENSIONS

Dimensions shown in inches and (mm).

14-Lead CERPAK (QC-14)



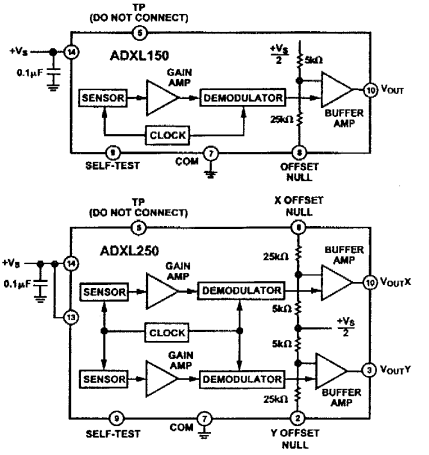
±5 *g* to±50 *g*, Low Noise, Low Power, Single/Dual Axis *i*MEMS® Accelerometers

ADXL150/ADXL250

FEATURES

Complete Acceleration Measurement System on a Single Monolithic IC
80 dB Dynamic Range
Pin Programmable ±50 g or ±25 *g* Full Scale
Low Noise: 1 mg/√Hz Typical
Low Power: <2 mA per Axis
Supply Voltages as Low as 4 V
2-Pole Filter On-Chip
Ratiometric Operation
Complete Mechanical & Electrical Self-Test
Dual & Single Axis Versions Available
Surface Mount Package

FUNCTIONAL BLOCK DIAGRAMS



GENERAL DESCRIPTION

The ADXL150 and ADXL250 are third generation ±50 *g* surface micromachined accelerometers. These improved replacements for the ADXL50 offer lower noise, wider dynamic range, reduced power consumption and improved zero *g* bias drift.

The ADXL150 is a single axis product; the ADXL250 is a fully integrated dual axis accelerometer with signal conditioning on a single monolithic IC, the first of its kind available on the commercial market. The two sensitive axes of the ADXL250 are orthogonal (90°) to each other. Both devices have their sensitive axes in the same plane as the silicon chip.

The ADXL150/ADXL250 offer lower noise and improved signal-to-noise ratio over the ADXL50. Typical S/N is 80 dB, allowing resolution of signals as low as 10 mg, yet still providing a ±50 *g* full-scale range. Device scale factor can be increased from 38 mV/*g* to 76 mV/*g* by connecting a jumper between V_{OUT} and the offset null pin. Zero *g* drift has been reduced to 0.4 *g* over the industrial temperature range, a 10× improvement over the ADXL50. Power consumption is a modest 1.8 mA per axis. The scale factor and zero *g* output level are both ratiometric to the power supply, eliminating the need for a voltage reference when driving ratiometric A/D converters such as those found in most microprocessors. A power supply bypass capacitor is the only external component needed for normal operation.

The ADXL150/ADXL250 are available in a hermetic 14-lead surface mount cerpac package specified over the 0°C to +70°C commercial and –40°C to +85°C industrial temperature ranges. Contact factory for availability of devices specified over automotive and military temperature ranges.

*i*MEMS is registered trademark of Analog Devices, Inc.

REV. 0

Information furnished by Analog Devices is believed to be accurate and reliable. However, no responsibility is assumed by Analog Devices for its use, nor for any infringements of patents or other rights of third parties which may result from its use. No license is granted by implication or otherwise under any patent or patent rights of Analog Devices.

One Technology Way, P.O. Box 9106 Norwood, MA 02062-9106, U.S.A.
Tel: 781/329-4700 World Wide Web Site: http://www.analog.com
Fax: 781/326-8703
© Analog Devices, Inc., 1998

ADXL150/ADXL250–SPECIFICATIONS

(T_A = +25°C for J Grade, T_A = 40°C to +85°C for A Grade, V_S = +5.00 V, Acceleration = Zero *g*, unless otherwise noted)

		ADXL150JQC/AQC			ADXL250JQC/AQC			
Parameter	**Condition**	**Min**	**Typ**	**Max**	**Min**	**Typ**	**Max**	**Units**
SENSOR								
Guaranteed Full-Scale Range		±40	±50		±40	±50		*g*
Nonlinearity			0.2			0.2		% of FS
Package Alignment Error[1]			±1			±1		Degrees
Sensor-to-Sensor Alignment Error						±0.1		Degrees
Transverse Sensitivity[2]			±2			±2		%
SENSITIVITY								
Sensitivity (Ratiometric)[3]	Y Channel				33.0	38.0	43.0	mV/*g*
	X Channel	33.0	38.0	43.0	33.0	38.0	43.0	mV/*g*
Sensitivity Drift Due to Temperature	Delta from 25°C to T_{MIN} or T_{MAX}		±0.5			±0.5		%
ZERO *g* BIAS LEVEL								
Output Bias voltage[4]		V_S/2–0.35	V_S/2	V_S/2+0.35	V_S/2–0.35	V_S/2	V_S/2+0.35	V
Zero *g* Drift Due to Temperature	Delta from 25°C to T_{MIN} or T_{MAX}		0.2			0.3		*g*
ZERO-*g* OFFSET ADJUSTMENT								
Voltage Gain	Delta V_{OUT}/Delta $V_{OS\ PIN}$	0.45	0.50	0.55	0.45	0.50	0.55	V/V
Input Impedence		20	30		20	30		kΩ
NOISE PERFORMANCE								
Noise Density[5]			1	2.5		1	2.5	$mg/\sqrt{Hz}$
Clock Noise			5			5		mV p-p
FREQUENCY RESPONSE								
–3 dB Bandwidth		900	1000		900	1000		Hz
Bandwidth Temperature Drift	T_{MIN} to T_{MAX}		50			50		kHz
Sensor Resonant Frequency	Q = 5		24			24		kHz
SELF-TEST								
Output Change	ST Pin from Logic "0" to '1"	0.25	0.40	0.60	0.25	0.40	0.60	V
Logic "1" Voltage		V_S–1			V_S–1			V
Logic "0" Voltage	To Common			1.0			1.0	V
Input Resistance		30	50		30	50		kΩ
OUTPUT AMPLIFIER								
Output Voltage Swing	I_{OUT} = ±100 µA	0.25		V_S–0.25	0.25		V_S–0.25	V
Capacitive Load Drive		1000			1000			pF
POWER SUPPLY (V_S)[7]								
Functional Voltage Range		4.0		6.0	4.0		6.0	V
Quiescent Supply Current	ADXL150		1.8	3.0				mA
	ADXL250 (Total 2 Channels)					3.5	5.0	mA
TEMPERATURE RANGE								
Operating Range J		0		+70	0		+70	°C
Specified Performance A		–40		+85	–40		+85	°C

NOTES

[1]Alignment error is specified as the scale between the ture axis of sensitivity and the edge of the package.

[2]Transverse sensitivity is measured with an applied acceleration that is 90 degrees from the indicated axis of sensitivity.

[3]Ratiometric: $V_{OUT} = V_S/2 + (\text{Sensitivity} \times V_S/5\ V \times a)$ where a = applied acceleration in *g*s, and V_S = supply voltage. See Figure 21. Output scale factor can be doubled by connecting V_{OUT} to the offset null pin.

[4]Ratiometric, proportional to V_S/2. See Figure 21.

[5]See Figure 11 and Device Bandwidth vs. Resolution section.

[6]Sclf-test output varies with supply voltage.

[7]When wing ADXL250, both Pins 13 and 14 must be connected to the supply for the device to function.

Specifications subject to change without notice.

ABSOLUTE MAXIMUM RATINGS*

Acceleration (Any Axis, Unpowered for 0.5 ms) . . . . . 2000 *g*
Acceleration (Any Axis, Powered for 0.5 ms) . . . . . 500 *g*
$+V_S$ . . . . . –0.3 V to –7.0 V
Output Short Circuit Duration
(V_{OUT}, V_{REF} Terminals to Common) . . . . . Indefinite
Operating Temperature . . . . . –55°C to +125ºC
Storage Temperature . . . . . –65°C to +150°C

*Stresses above those listed under Absolute Maximum Ratings may cause permanent damage to the device. This is a stress rating only; the functional operation of the device at these or any other conditions above those indicated in the operational sections of this specification is not implied. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

Drops onto hard surfaces can cause shocks of greater than 2000 *g* and exceed the absolute maximum rating of the device. Care should be exercised in handling to avoid damage.

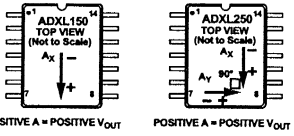


Figure 1. ADXL150 and ADXL250 Sensitive Axis Orientation

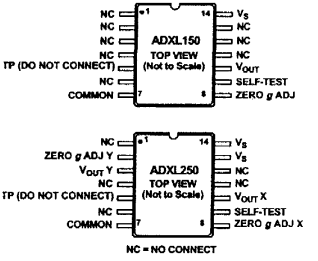
Package Characteristics

Package	θ_{JA}	θ_{JC}	Device Weight
14-Lead CERPAK	110°C/W	30°C/W	5 Grams

ORDERING GUIDE

Model	Temperature Range
ADXL150JQC	0°C to +70°C
ADXL150AQC	–40°C to +85°C
ADXL250JQC	0°C to +70°C
ADXL250AQC	–40°C to +85°C

PIN CONNECTIONS



NOTE: WHEN USING ADXL250, BOTH PINS 13 AND 14 NEED TO BE CONNECTED TO SUPPLY FOR DEVICE TO FUNCTION

CAUTION

ESD (electrostatic discharge) sensitive device. Electrostatic charges as high as 4000 V readily accumulate on the human body and test equipment and can discharge without detection. Although the ADXL150/ADXL250 features proprietary ESD protection circuitry, permanent damage may occur on devices subjected to high energy electrostatic discharges. Therefore, proper ESD precautions are recommended to avoid performance degradation or loss of functionality.

GLOSSARY OF TERMS

Acceleration: Change in velocity per unit time.

Acceleration Vector: Vector describing the net acceleration acting upon the ADXL150/ADXL250.

***g*:** A unit of acceleration equal to the average force of gravity occurring at the earth's surface. A g is approximately equal to 32.17 feet/s^2 or 9.807 meters/s^2.

Nonlinearity: The maximum deviation of the ADXL150/ADXL250 output voltage from a best fit straight line fitted to a plot of acceleration vs. output voltage, calculated as a % of the full-scale output voltage (at 50 *g*).

Resonant Frequency: The natural frequency of vibration of the ADXL150/ADXL250 sensor's central plate (or "beam"). At its resonant frequency of 24 kHz, the ADXL150/ADXL250's moving center plate has a slight peak in its frequency response.

Sensitivity: The output voltage change per g unit of acceleration applied, specified at the V_{OUT} pin in mV/*g*.

Total Alignment Error: Net misalignment of the ADXL150/ADXL250's on-chip sensor and the measurement axis of the application. This error includes error due to sensor die alignment to the package, and any misalignment due to installation of the sensor package in a circuit board or module.

Transverse Acceleration: Any acceleration applied 90° to the axis of sensitivity.

Transverse Sensitivity Error: Ile percent of a transverse acceleration that appears at V_{OUT}.

Transverse Axis: The axis perpendicular (90°) to the axis of sensitivity.

Zero *g* Bias Level: The output voltage of the ADXL150/ADXL250 when there is no acceleration (or gravity) acting upon the axis of sensitivity. The output offset is the difference between the actual zero *g* bias level and (V_S/2).

Polarity of the Acceleration Output

The polarity of the ADXL150/ADXL250 output is shown in Figure 1. When its sensitive axis is oriented to the earth's gravity (and held in place), it will experience in acceleration of +1 *g*. This corresponds to a change of approximately +38 mV at the output pin. Note that the polarity will be reversed if the package is rotated 180º. The figure shows the ADXL250 oriented so that its "X" axis measures +1 *g*. If the package is rotated 90º clockwise (Pin 14 up, Pin 1 down), the ADXL250's "Y" axis will now measure +1 *g*.

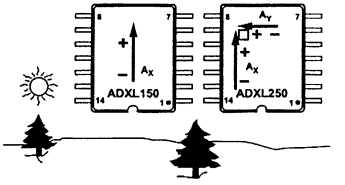


Figure 2. Output Polarity

Acceleration Vectors

The ADXL150/ADXL250 is a sensor designed to measure accelerations that result from an applied force. It responds to the component of acceleration on its sensitive X axis (ADXL150) or on both the "X" and "Y" axis (ADXL250).

Typical Characteristics (@+5 V dc, +25°C with a 38 mV/*g* Scale Factor unless otherwise noted)

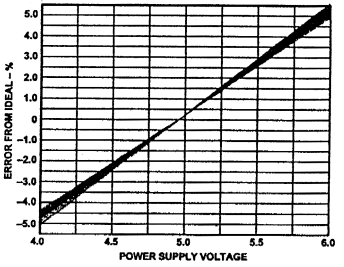


Figure 3. Typical Sensitivity Error from Ideal Ratiometric Response for a Number of Units

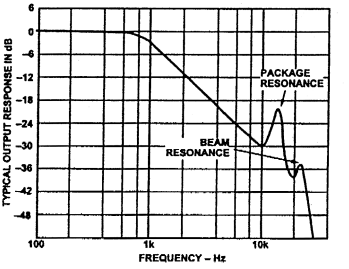


Figure 6. Typical Output Response vs. Frequency of ADXL150/ADXL250 on a PC Board that Has Been Conformally Coated

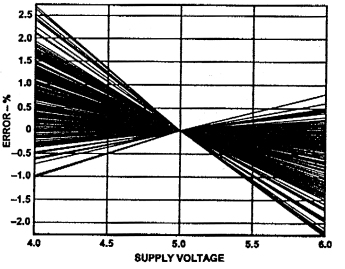


Figure 4. Offset Error of Zero g Level from Ideal $V_S/2$ Response as a Percent of Full-Scale for a Number of Units

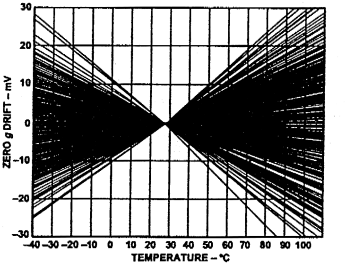


Figure 7. Typical Zero g Drift for a Number of Units

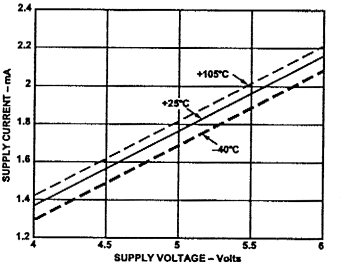


Figure 5. Typical Supply Current vs. Supply Voltage

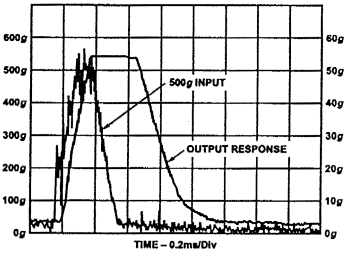


Figure 8. Typical 500 g Step Recovery at the Output

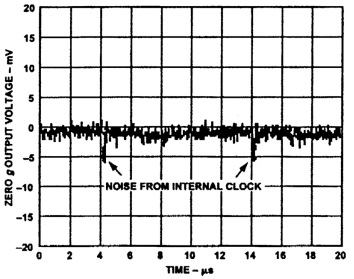


Figure 9. Typical Output Noise Voltage with Spikes Generated by Internal Clock

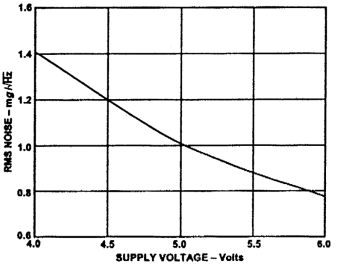


Figure 12. Noise vs. Supply Voltage

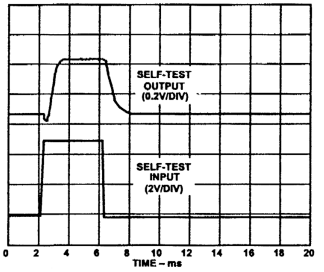


Figure 10. Typical Self-Test Response

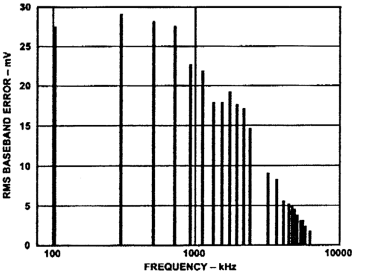


Figure 13. Baseband Error Graph

Figure 13 shows the mV rms error in the output signal if there is a noise on the power supply pin of 1 mV rms at the internal clock frequency or its odd harmonics. This is a baseband noise and can be at any frequency in the 1 kHz passband or at dc.

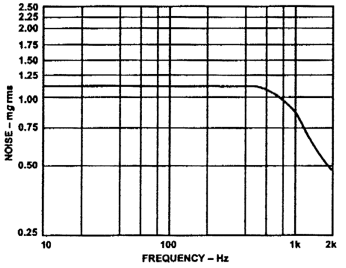


Figure 11. Noise Spectral Density

THEORY OF OPERATION

The ADXL150 and ADXL250 are fabricated using a proprietary surface micromachining process that has been in high volume production since 1993. The fabrication technique uses standard integrated circuit manufacturing methods enabling all the signal processing circuitry to be combined on the same chip with the sensor.

The surface micromachined sensor element is made by depositing polysilicon on a sacrificial oxide layer that is then etched away leaving the suspended sensor element. Figure 14 is a simplified view of the sensor structure. The actual sensor has 42 unit cells for sensing acceleration. The differential capacitor sensor is composed of fixed plates and moving plates attached to the beam that moves in response to acceleration. Movement of the beam changes the differential capacitance, which is measured by the on chip circuitry.

The sensor has 12-unit capacitance cells for electrostatically forcing the beam during a self-test. Self-test is activated by the user with a logic high on the self-test input pin. During a logic high, an electrostatic force acts on the beam equivalent to approximately 20% of full-scale acceleration input, and thus a proportional voltage change appears on the output pin. When activated, the self-test feature exercises both the entire mechanical structure and the electrical circuitry.

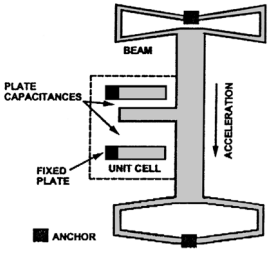


Figure 14. Simplified View of Sensor Under Acceleration

All the circuitry needed to drive the sensor and convert the capacitance change to voltage is incorporated on the chip requiring no external components except for standard power supply decoupling. Both sensitivity and the zero-g value are ratiometric to the supply voltage, so that ratiometeric devices following the accelerometer (such as an ADC, etc.) will track the accelerometer if the supply voltage changes. The output voltage (V_{OUT}) is a function of both the acceleration input (a) and the power supply voltage (V_S) as follows:

$$V_{OUT} = V_S/2 - (Sensitivity \times \frac{V_S}{5V} \times a)$$

Both the ADXL150 and ADYCL250 have a 2-pole Bessel switched-capacitor filter. Bessel filters, sometimes called linear phase filters, have a step response with minimal overshoot and a maximally flat group delay. The –3 dB frequency of the poles is preset at the factory to 1 kHz. These filters are also completely self-contained and buffered, requiring no external components.

MEASURING ACCELERATIONS LESS THAN 50 *g*

The ADXL150/ADXL250 require only a power supply bypass capacitor to measure ±50 g accelerations. For measuring ±50 *g* accelerations, the accelerometer may be directly connected to an ADC (see Figure 25). The device may also be easily modified to measure lower *g* signals by increasing its output wale factor.

The scale factor of an accelerometer specifies the voltage change of the output per *g* of applied acceleration. This should not be confused with its resolution. The resolution of the device is the lowest *g* level the accelerometer is capable of measuring. Resolution is principally determined by the device noise and the measurement bandwidth.

The zero *g* bias level is simply the dc output Level of the accelerometer when it is not in motion or being acted upon by the earth's gravity.

Pin Programmable Scale Factor Option

In its normal state, the ADXL150/ADXL250's buffer amplifier provides in output scale factor of 38 mV/*g*, which is set by an internal voltage divider. This gives a full-scale range of +50 *g* and a nominal bandwidth of 1 kHz.

A factor-of-two increase in sensitivity can be obtained by connecting the V_{OUT} pin to the offset null pin, assuming that it is not needed for offset adjustment. This connection has the effect of reducing the internal feedback by a factor of two, doubling the buffer's gain. This increases the output scale factor to 76 mV/*g* and provides a ±25 g full-scale range.

Simultaneously, connecting these two pins also increases the amount of internal post filtering, reducing the noise floor and changing the nominal 3 dB bandwidth of the ADXL150/ADXL250 to 500 Hz. Note that the post filter's "Q" will also be reduced by a factor of $\sqrt{2}$ from 0.58 (Bessel response) to a much gentler "Q" value of 0.41. The primary effect of this change in "Q" is only at frequencies within two octaves of the corner frequency; above this the two filter slopes am essentially the same. In applications where a flat response up to 500 Hz is needed, it is better to operate the device at 38 mV/*g* and use an external post filter. Note also that connecting V_{OUT} to the offset pin adds a 30 kΩ load from V_{OUT} to $V_S/2$. When swinging ±2 V at V_{OUT}, this added load will consume ±60 µA of the ADXL150/ADXL250's 100 µA (typical) output current drive.

Increasing the *i*MEMS Accelerometer's Output Scale Factor

Figure 15 shows the basic connections for using an external buffer amplifier to increase die output scale factor.

The output multiplied by the gain of the buffer, which is simply the value of resistor R3 divided by RI. Choose a convenient scale factor, keeping in mind that the buffer pin not only amplifies the signal, but my noise or drift as well. Too much pin can also cause the buffer to saturate and clip the output waveform.

Note that the "+" input of the external op amp uses the offset null pin of the ADXL150/ADXL250 as a reference, biasing the op amp at midsupply, saving two resistors and reducing power consumption. The offset null pin connects to the $V_S/2$ reference point inside the accelerometer via 30 kΩ, so it is important not to load this pin with more dim a few microamps.

It is important to use a single-supply or "rail-to-rail" op amp for the external buffer as it needs to be able to swing close to the supply and ground.

The circuit of Figure 15 is entirely adequate for many applications, but its accuracy is dependent on the pretrimmed accuracy of the accelerometer and this will vary by product type and grade. For the highest possible accuracy, an external trim is mended. As shown by Figure 20, this consists of a potentiometer Rla, in series with a fixed resistor, Rlb. Another to select resistor values after measuring the device's scale (see Figure 17).

AC Coupling

If a dc (gravity) response is not required—for example ** tion measurement applications—ac coupling can be ** between the accelerometer's output and the external op** input as shown in Figure 16. The use of ac coupling ** eliminates my zero g drift and allows the maximum ** amp gain without clipping.

Resistor R2 and capacitor C3 together form a high ** whose corner frequency is 1/(2 x R2 C3). This filter ** the signal from the accelerometer by 3 dB at the **, and it will continue to reduce it at a rate of 6 ** (20 dB per decade) for signals below the corner frequ ** Capacitor CBS should be a nonpolarized, low leakage type **

If ac coupling is used, the self-test feature must be ** the accelerometer's output rather than at the external ** output (since the self-test output is a dc voltage).

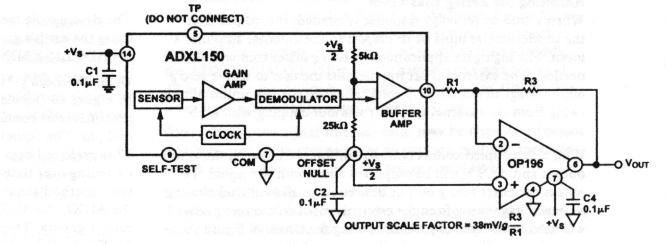
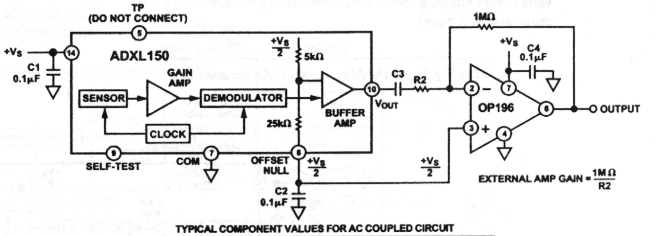


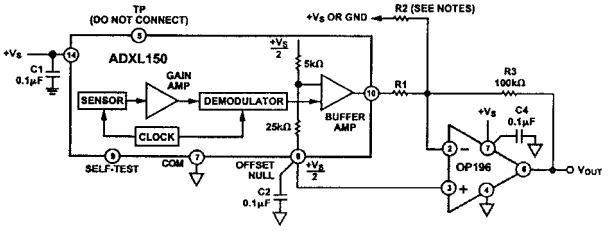
Figure 15. Using an External Op Amp to Increase Output Scale Factor



TYPICAL COMPONENT VALUES FOR AC COUPLED CIRCUIT

BUFFER GAIN	FS RANGE	R2	C3 VALUE FOR 3dB CORNER FREQ 1Hz	3Hz	10Hz	20Hz
2	±25g	1MΩ	0.15µF	0.05µF	0.015µF	0.0075µF
4	±12.5g	332kΩ	0.47µF	0.15µF	0.047µF	0.022µF
5	±10g	249kΩ	0.68µF	0.22µF	0.022µF	0.01µF

Figure 16. AC Coupled Connection Using an External Op Amp



NOTES:
0g "QUICK" CALIBRATION METHOD USING RESISTOR R2 AND A +5V SUPPLY.
(a) WITH ACCELEROMETER ORIENTED AWAY FROM EARTH'S GRAVITY (i.e., SIDEWAYS), MEASURE PIN 10 OF THE ADXL150.
(b) CALCULATE THE OFFSET VOLTAGE THAT NEEDS TO BE NULLED: $V_{OS} = (+2.5V - V_{PIN}(10)(R3/R1)$.
(c) $R2 = \frac{2.5\ V\ (R3)}{V_{OS}}$
(d) FOR V_{PIN} 10 > +2.5V, R2 CONNECTS TO GND.
(e) FOR V_{PIN} 10 < +2.5V, R2 CONNECTS TO $+V_S$.

DESIRED OUTPUT SCALE FACTOR	FS RANGE	EXT AMP GAIN	R1 VALUE
76mV/*g*	±25*g*	2.0	49.9kΩ
100mV/*g*	±20*g*	2.6	38.3kΩ
200mV/*g*	±10*g*	5.3	18.7kΩ
400mV/*g*	±5*g*	10.5	9.53kΩ

Figure 17. "Quick Zero g Calibration" Connection

Adjusting the Zero *g* Bias Level

When a true dc (gravity) response is needed, the output from the accelerometer must be dc coupled to the external amplifier's input. For high gain applications, a zero *g* offset trim will also be needed. The external offset trim permits the user to set the zew *g* offset voltage to exactly +2.5 volts (allowing the maximum output swing from the external amplifier without clipping with a +5 supply).

With a dc coupled connection, any difference between the zero *g* output and +2.5 V will be amplified along with the signal. To obtain the exact zero *g* output desired or to allow the maximum output voltage swing from the external amplifier, the zero *g* offset will need to be externally trimmed using the circuit of Figure 20.

The external amplifier's maximum output swing should be limited to ±2 volts, which provides a safety margin of ±0.25 volts before clipping. With a +2.5 volt zero *g* level, the maximum gain will equal:

$$\frac{2\ Volts}{38\ mV/g\ Times\ the\ Max\ Applied\ Acceleration\ in\ g}$$

The device scale factor and zero g offset levels can be calibrated using the earth's gravity, as explained in the section "calibrating the ADXL150/ADXL250."

Using the Zero *g* "Quick-Cal" Method

In Figure 18 (accelerometer alone, no external op amp), a trim potentiometer connects directly to the accelerometer's zero *g* null pin. The "quick offset calibration" scheme shown in Figure 17 is preferred over using a potentiometer, which could change its setting over time due to vibration. The "quick offset calibration" method requires measuring only the output voltage of the ADXL150/ADXL250 while it is oriented normal to the earth's gravity. Then, by using the simple equations shown in the figures, the correct resistance value for R2 can be calculated. In Figure 17, an external op amp is used to amplify the signal. A resistor, R2, is connected to the op amp's summing junction. The other side of R2 connects to either ground or $+V_S$ depending on which direction the offset needs to be shifted.

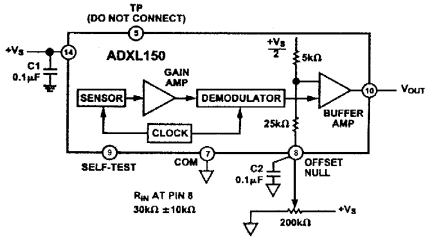


Figure 18. Offset Nulling the ADXL150/ADXL250 Using a Trim Potentiometer

DEVICE BANDWIDTH VS. MEASUREMENT RESOLUTION

Although an accelerometer is usually specified according to its full-scale *g* level, the limiting resolution of the device, i.e., its minimum discernible input level, is extremely important when measuring low *g* accelerations.

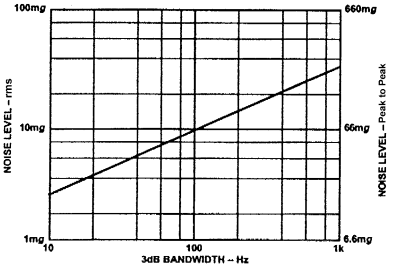


Figure 19. ADXL150/ADXL250 Noise Level vs. 3 dB Bandwidth (Using a "Brickwall" Filter)

The limiting resolution is predominantly set by the measurement noise "floor," which includes the ambient background noise and the noise of the ADXL150/ADXL250 itself. The level of the noise floor varies directly with the bandwidth of the measurement. As the measurement bandwidth is reduced, the noise floor drops, improving the signal-to-noise ratio of the measurement and increasing its resolution.

The bandwidth of the accelerometer can be easily reduced by adding low-pass or bandpass filtering. Figure 19 shows the typical noise vs. bandwidth characteristic of the ADXL150/ADXL250.

The output noise of the ADXL150/ADXL250 scales with the square root of the measurement bandwidth. With a single pole roll-off, the equivalent rms noise bandwidth is π divided by 2 or approximately 1.6 times the 3 dB bandwidth. For example, the typical rms noise of the ADXL150 using a 100 Hz one pole post filter is:

$$Noise(rms) = 1\ mg/\sqrt{Hz} \times \sqrt{100(1.6)} = 12.25\ mg$$

Because the ADXL150/ADXL250's noise is, for all practical purposes, Gaussian in amplitude distribution, the highest noise amplitudes have die smallest (yet nonzero) probability. Peak-to-peak noise is therefore difficult to measure and can only be estimated due to its statistical nature. Table I is useful for estimating the probabilities of exceeding various peak values, given the rms value.

TABLE I.

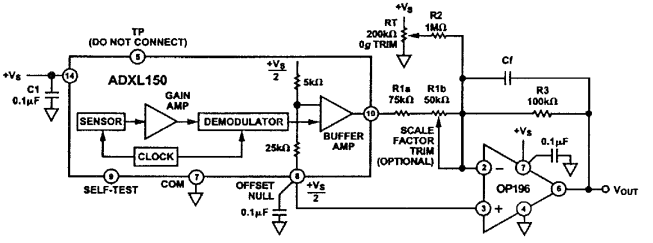
Nominal Peak-to-Peak Value	% of Time that Noise Will Exceed Nominal Peak-to-Peak Value
2.0 × rms	32%
4.0 × rms	4.6%
6.0 × rms	0.27%
6.6 × rms	0.1%
8.0 × rms	0.006%

RMS and peak-to-peak noise (for 0. 1% uncertainty) for various bandwidths are estimated in Figure 19. As shown by the figure, device noise drops dramatically as the operating bandwidth is reduced. For example, when operated in a 1 kHz bandwidth, the ADXL150/ADXL250 typically have an rms noise level of 32 mg. When the device bandwidth is rolled off to 100 Hz, the noise level is reduced to approximately 10 mg.

Alternatively, the signal-to-noise ratio may be improved considerably by using a microprocessor to perform multiple measurements and then to compute the average signal level.

Low-Pass Filtering

The bandwidth of the accelerometer can easily be reduced by using post filtering. Figure 20 shows how the buffer amplifier can be connected to provide 1-pole post filtering, zero *g* offset trimming, and output scaling. The table provides practical component values



DESIRED OUTPUT SCALE FACTOR	F.S. RANGE	EXT AMP GAIN	R3 VALUE	Cf (µF) 100Hz	Cf (µF) 30Hz	Cf (µF) 10Hz
76m/*g*	±25*g*	2.0	200kΩ	0.0082	0.027	0.082
100m/*g*	±20*g*	2.6	261kΩ	0.0056	0.022	0.056
200m/*g*	±10*g*	5.3	536kΩ	0.0033	0.010	0.033
400m/*g*	±5*g*	10.5	1MΩ	0.0015	0.0056	0.015

Figure 20. One-Pole Post Filter Circuit with SF and Zero g Offset Trims

for various full-scale *g* levels and approximate circuit bandwidths. For bandwidths other than those listed, use the formula:

$$Cf = \frac{1}{(2\pi R3)\ Desired\ 3dB\ Bandwidth\ in\ Hz}$$

or simply scale the value of capacitor Cf accordingly; i.e., for an application with a 50 Hz bandwidth, the value of Cf will need to be twice as large as its 100 Hz value. If further noise reduction is needed while maintaining the maximum possible bandwidth, a 2- or 3-pole post filter is recommended. These provide a much steeper roll-off of noise above the pole frequency. Figure 21 shows a circuit that provides 2-pole post filtering. Component values for the 2-pole filter were selected to operate the first op amp at unity gain. Capacitors C3 and C4 were chosen to provide 3 dB bandwidths of 10 Hz, 30 Hz, 100 Hz and 300 Hz.

The second op amp offsets and scales the output to provide a +2.5 V ± 2 V output over a wide range of full-scale *g* levels.

APPLICATION HINTS

ADXL250 Power Supply Pins

When wiring the ADXL250, be sure to connect BOTH power supply terminals, Pins 14 and 13.

Ratiometric Operation

Ratiometric operation means that the circuit uses the power supply as its voltage reference. If the supply voltage varies, the accelerometer and the other circuit components (such as an ADC, etc.) track each other and compensate for the change.

Figure 22 shows how both the zero *g* offset and output sensitivity of the ADXL150/ADXL250 vary with changes in supply voltage. If they are to be used with nonratiometric devices, such as an ADC with a built-in 5 V reference, then both components should be referenced to the same source, in this case the ADC reference. Alternatively, the circuit can be powered from an external +5 volt reference.

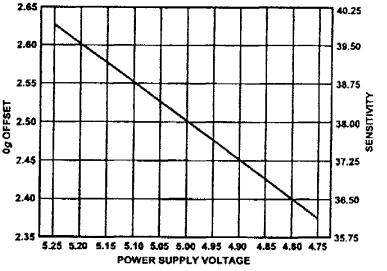
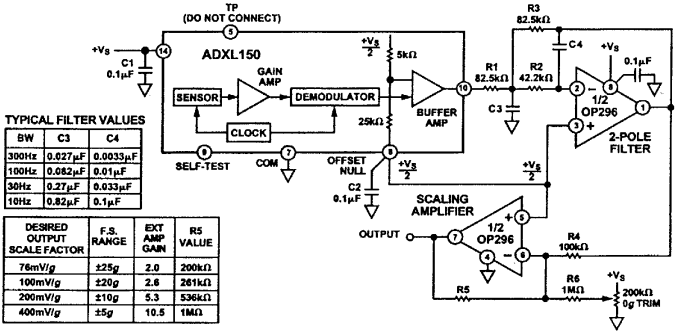


Figure 22. Typical Ratiornetric Operation

Since any voltage variation is transferred to the accelerometer's output, it is important to reduce any power supply noise. Simply following good engineering practice of bypassing the power supply right at Pin 14 of the ADXL150/ADXL250 with a 0.1 µF capacitor should be sufficient.



TYPICAL FILTER VALUES

BW	C3	C4
300Hz	0.027µF	0.0033µF
100Hz	0.082µF	0.01µF
30Hz	0.27µF	0.033µF
10Hz	0.82µF	0.1µF

DESIRED OUTPUT SCALE FACTOR	F.S. RANGE	EXT AMP GAIN	R5 VALUE
76mV/*g*	±25*g*	2.0	200kΩ
100mV/*g*	±20*g*	2.6	261kΩ
200mV/*g*	±10*g*	5.3	536kΩ
400mV/*g*	±5*g*	10.5	1MΩ

Figure 21. Two-Pole Post Filter Circuit

Additional Noise Reduction Techniques

Shielded wire should be used for connecting the accelerometer to any circuitry that is more than a few inches away—to avoid 60 Hz pickup from ac line voltage. Ground the cable's shield at only one end and connect a separate common lead between the circuits; this will help to prevent ground loops. Also, if the accelerometer is inside a metal enclosure, this should be grounded as well.

Mounting Fixture Resonances

A common source of error in acceleration sensing is resonance of the mounting fixture. For example, the circuit board that the ADXL150/ADXL250 mounts to may have resonant frequencies in the same range as the signals of interest. This could cause the signals measured to be larger than they really are. A common solution to this problem is to damp these resonances by mounting the ADXL150/ADXL250 near a mounting post or by adding extra screws to hold the board more securely in place.

When testing the accelerometer in your end application, it is recommended that you test the application at a variety of frequencies to ensure that no major resonance problems exist.

REDUCING POWER CONSUMPTION

The use of a simple power cycling circuit provides a dramatic reduction in the accelerometer's average current consumption. In low bandwidth applications such as shipping recorders, a simple, low cost circuit can provide substantial power reduction.

If a microprocessor is available, it can supply a TTL clock pulse to toggle the accelerometer's power on and off.

A 10% duty cycle, 1 ms on, 9 ms off, reduces the average current consumption of the accelerometer from 1.8 mA to 180 µA, providing a power reduction of 90%.

Figure 23 shows the typical power-on settling time of the ADXL150/ADXL250.

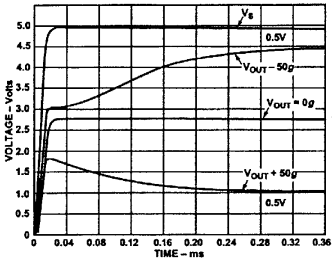


Figure 23. Typical Power-On Settling with Full-Scale Input. Time Constant of Post Filter Dominates the Response When a Signal Is Present.

CALIBRATING THE ADXL150/ADXL250

If a calibrated shaker is not available, both the zero *g* level and scale factor of the ADXL150/ADXL250 may be easily set to fair accuracy by using a self-calibration technique based on the 1 *g* acceleration of the earth's gravity. Figure 24 shows how gravity and package orientation affect the ADXL150/ADXL250's output. With its axis of sensitivity in the vertical plane, the ADXL150/ADXL250 should register a 1 *g* acceleration, either positive or negative, depending on orientation. With the axis of sensitivity in the horizontal plane, no acceleration (the zero *g* bias level) should be indicated. The use of an external buffer amplifier may invert the polarity of the signal.

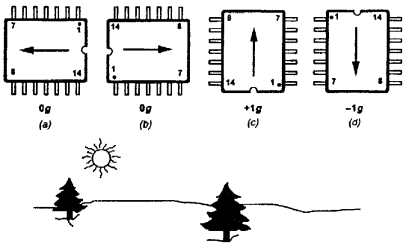


Figure 24. Using the Earth's Gravity to Self-Calibrate the ADXL150/ADXL250

Figure 24 shows how to self-calibrate the ADXL150/ADXL250. Place the accelerometer on its side with its axis of sensitivity oriented as shown in "a." (For the ADXL250 this would be the "X" axis—its "Y" axis is calibrated in the same manner, but the part is rotated 90° clockwise.) The zero *g* offset potentiometer RT is then roughly adjusted for midscale: +2.5 V at the external amp output (see Figure 20).

Next, the package axis should be oriented as in "c" (pointing down) and the output reading noted. The package axis should then be rotated 180° to position "d" and the scale factor potentiometer, Rlb, adjusted so that the output voltage indicates a change of 2 *g*s in acceleration. For example, if the circuit scale factor at the external buffer's output is 100 mV per *g*, the scale factor trim should be adjusted so that an output change of 200 mV is indicated.

Self-Test Function

A Logic "1" applied to the self-test (ST) input will cause an electrostatic force to be applied to the sensor that will cause it to deflect. If the accelerometer is experiencing an acceleration when the self-test is initiated, the output will equal the algebraic sum of the two inputs. The output will stay at the self-test level as long as the ST input remains high, and will return to the actual acceleration level when the ST voltage is removed.

Using an external amplifier to increase output scale factor may cause the self-test output to overdrive the buffer into saturation. The self-test may still be used in this case, but the change in the output must then be monitored at the accelerometer's output instead of the external amplifier's output.

Note that the value of the self-test delta is not an exact indication of the sensitivity (mV/*g*) and therefore may not be used to calibrate the device for sensitivity error.

MINIMIZING EMI/RFI

The architecture of the ADXL150/ADXL250, and its use of synchronous demodulation, makes the device immune to most electromagnetic (EMI) and radio frequency (RFI) interference. The use of synchronous demodulation allows the circuit to reject all signals except those at the frequency of the oscillator driving the sensor element. However, the ADXL150/ADXL250 have a sensitivity to noise on the supply lines that is near its internal clock frequency (approximately 100 kHz) or its odd harmonics and can exhibit baseband errors at the output. These error signals are the beat frequency signals between the clock and the supply noise.

Such noise can be generated by digital switching elsewhere in the system and must be attenuated by proper bypassing. By inserting a small value resistor between the accelerometer and its power supply, an RC filter is created. This consists of the resistor and the accelerometer's normal 0.1 µF bypass capacitor. For example if R = 20 Ω and C = 0.1 µF, a filter with a pole at 80 kHz is created, which is adequate to attenuate noise on the supply from most digital circuits, with proper ground and supply layout.

Power supply decoupling, short component leads, physically small (surface mount, etc.) components and attention to good grounding practices all help to prevent RFI and EMI problems. Good grounding practices include having separate analog and digital grounds (as well as separate power supplies or very good decoupling) on the printed circuit boards.

INTERFACING THE ADXL150/ADXL250 SERIES *i*MEMS ACCELEROMETERS WITH POPULAR ANALOG-TO-DIGITAL CONVERTERS.

Basic Issues

The ADXL150/ADXL250 Series accelerometers were designed to drive popular analog-to-digital converters (ADCs) directly. In applications where both a ±50 *g* full-scale measurement range and a 1 kHz bandwidth are needed, the V_{OUT} terminal of the accelerometer is simply connected to the V_{IN} terminal of the ADC as shown in Figure 25a. The accelerometer provides its (nominal) factory preset scale factor of +2.5 V ±38 mV/*g* which drives the ADC input with +2.5 V ±1.9 V when measuring a 50 *g* full-scale signal (38 mV/*g* × 50 *g* = 1.9 V).

As stated earlier, the use of post filtering will dramatically improve the accelerometer's low *g* resolution. Figure 25b shows a simple post filter connected between the accelerometer and the ADC. This connection, although easy to implement, will require fairly large values of Cf, and the accelerometer's signal will be loaded down (causing a scale factor error) unless the ADC's input impedance is much greater than the value of Rf. ADC input impedance's range from less than 1.5 kΩ up to greater than 15 kΩ with 5 kΩ values being typical. Figure 25c is the preferred connection for implementing low-pass filtering with the added advantage of providing an increase in scale factor, if desired.

Calculating ADC Requirements

The resolution of commercial ADCs is specified in bits. In an ADC, the available resolution equals 2^n, where *n* is the number of bits. For example, an 8-bit converter provides a resolution of 2^8 which equals 256. So the full-scale input range of the converter divided by 256 will equal the smallest signal it can resolve.

In selecting an appropriate ADC to use with our accelerometer we need to find a device that has a resolution better than the measurement resolution but, for economy's sake, not a great deal better.

For most applications, an 8- or 10-bit converter is appropriate. The decision to use a 10-bit converter alone, or to use a gain stage together with an 8-bit converter, depends on which is more important: component cost or parts count and ease of assembly,

Table II shows some of the tradeoffs involved.

Table II.

	8-Bit Converter and Op Amp Preamp	10-bit (or 12-Bit) Converter
Advantages:	Low Cost Converter	No Zero *g* Trim Required
Disadvantages:	Needs Op Amp Needs Zero *g* Trim	Higher Cost Converter

Adding amplification between the accelerometer and the ADC will reduce the circuit's full-scale input range but will greatly reduce the resolution requirements (and therefore the cost) of the ADC. For example, using an op amp with a gain of 5.3 following the accelerometer will increase the input drive to the ADC from 38 mV/*g* to 200 mV/*g*. Since the signal has been gained up, but the maximum full-scale (clipping) level is still the same, the dynamic range of the measurement has also been reduced by 5.3.

Table III. Typical System Resolution Using Some Popular ADCs Being Driven with and without an Op Amp Preamp

Converter Type	2^n	Converter mV/Bit (5 V/2^n)	Preamp Gain	SF in mV/*g*	FS Range in *g*'s	System Resolution in *g*'s (p-p)
8 Bit	256	19.5 mV	None	38	±50	0.51
	256	19.5 mV	2	76	±25	0.26
	256	19.5 mV	2.63	100	±20	0.20
	256	19.5 mV	5.26	200	±10	0.10
10 Bit	1,024	4.9 mV	None	38	±50	0.13
	1,024	4.9 mV	2	76	±25	0.06
	1,024	4.9 mV	2.63	100	±20	0.05
	1,024	4.9 mV	5.26	200	±10	0.02
12 Bit	4,096	1.2 mV	None	38	±50	0.03
	4,096	1.2 mV	2	76	±25	0.02
	4,096	1.2 mV	2.63	100	±20	0.01
	4,096	1.2 mV	5.26	200	±10	0.006

Table III is a chart showing the required ADC resolution vs. the scale factor of the accelerometer with or without a gain amplifier. Note that the system resolution specified in the table refers

to that provided by the converter and preamp (if used). It is necessary to use sufficient post filtering with the accelerometer to reduce its noise floor to allow full use of the converter's resolution (see post filtering section).

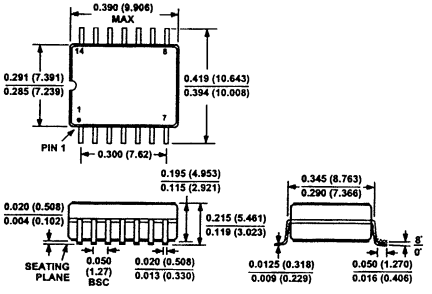
The use of a pin stage following the accelerometer will normally require the user to adjust the zero *g* offset level (either by trimming or by resistor selection—see previous sections).

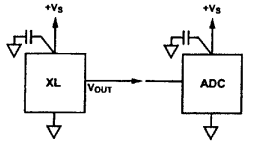
For many applications, a modern "economy priced" 10-bit converter, such as the AD7810 allows you to have high resolution without using a preamp or adding much to the overall circuit cost. In addition to simplicity and cost, it also meets two other necessary requirements: it operates from a single +5 V supply and is very low power.

OUTLINE DIMENSIONS

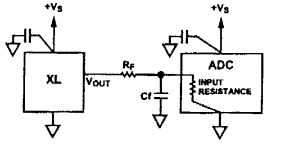
Dimensions shows in inches and (mm).

14-Lead Cerpac
(QC-14)

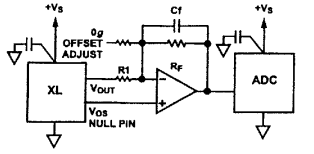




a. Direct Connection, No Signal Amplification or Post Filtering



b. Single-Pole Post Filtering, No Signal Amplification



c. Single-Pole Post Filtering and Signal Amplification

Figure 25. Interfacing the ADXL150/ADXL250 Series Accelerometers to an ADC

3.2. STRUCTURAL SYNTHESIS OF NANO- AND MICROELECTROMECHANICAL ACTUATORS AND SENSORS

New advances in micromachining and microstructures, nano- and microscale electromechanical devices, analog and digital ICs, provide enabling benefits and capabilities to design and manufacture NEMS and MEMS. Critical issues are to improve power and thermal management, circuitry and actuator/sensor integration, as well as embedded electronically controlled actuator/sensor assemblies. Very large scale integrated circuit and micromachining silicon, germanium, and gallium arsenic technologies have been developed and used to manufacture ICs and motion microstructures (microscale actuators and sensors). While enabling technologies have been developed to manufacture NEMS and MEMS, a spectrum of challenging problems remains. Electromagnetics and fluid dynamic, quantum phenomena, electro-thermo-mechanics and optics, biophysics and biochemistry, mechanical and structural synthesis, analysis and optimization, simulation and virtual prototyping, among other important problems, must be thoroughly studied in nano- and microscale. There are several key focus areas to be studied. In particular, structural synthesis and optimization, fabrication, nonlinear model development and analysis, system design and simulations.

An important problem addressed and studied in this section is the structural synthesis of motion nano- and microstructures (shape/geometry synthesis, optimization, and database developments). The proposed concept allows the designer to generate optimal structures of actuators and sensors. Using the proposed concept one can generate and optimize different nano- and microdevices, perform modeling and simulations, etc. These directly leverage high-fidelity model development and structural synthesis, allowing the designer to attain physical and behavioral (steady-state and transient) analysis, optimization, performance assessment, outcome prediction, etc.

3.2.1. Configurations and Structural Synthesis of Motion Nano- and Microstructures (Actuators and Sensors)

Using the structural synthesis concept, nano-, micro-, and miniscale actuators and sensors can be synthesized, analyzed, and optimized. In particular, electromechanical/electromagnetic-based motion nano- and microstructures (actuators and sensors) are classified using the specific classifiers, and the structural synthesis can be performed based upon different possible configurations, operating principles, phenomena, and physical laws.

We use the following electromagnetic systems

endless (E),

open-ended (O),

integrated (I),

and actuator/sensor geometry

plate (P),
spherical (S),
torroidal (T),
conical (N),
cylindrical (C) and
asymmetrical (A).

Using the possible electromagnetic systems and geometry, actuators and sensors (motion nano- and microstructures as well as nano- and microdevices) can be classified. From optimal structural and performance optimization viewpoints, a great number of factors influence the synthesis. For example, the designer must decide either rotational or translational devices (actuators and sensors) should be used. From electromagnetic standpoints, it is obvious that rotational-type actuators have higher power, torque and force densities, superior efficiency and performance compared with translational actuators. However, in many applications the translational actuators and sensors should be used due to specific kinematics, requirements, volume available, or size (for example, to measure the loads in the aircraft structures, the application of the translational-type sensor obviously is the preferable solution). Different rotational and translational actuators, as applied to displace the control surfaces and to change the geometry of flight surfaces, were illustrated in Figure 1.4.3 and 2.1.5.

The structural synthesis (geometrical design) and performance optimization of actuators and sensors is based on the consideration of the *electromotive* and *magnetomotive* forces, electromagnetic fields and magnetic structure, excitation, air gap and winding configurations, cooling and other important quantities. All possible actuator and sensor configurations can be classified using *endless* (closed) and *open-ended* (open) electromagnetic systems. This idea is extremely useful in studying existing and synthesizing novel motion structures and devices because an infinite number of innovative actuators and sensors can be synthesized. The application of a classifier is a starting point in structural synthesis and performance optimization. Advanced configurations can be synthesized and straightforwardly interpreted even without comprehensive analytical and numerical analysis (thorough analysis is needed as different actuator and sensor configurations are synthesized, and nonlinear electromagnetic-mechanical-thermal analysis must be performed and validated). It is evident that actuator/sensor geometry and electromagnetic systems play a central role in efficiency and performance. In fact, a motion structure (actuator/sensor) classifier, which is documented in Table 2.5.1, in addition to being qualitative, leads to quantitative analysis. Applying the cornerstone laws, the differential equations to model the actuator/sensor dynamics are straightforwardly derived, and analysis and simulations can be performed.

The actuators/sensors geometry and electromagnetic systems, as documented in Table 3.2.1, are partitioned into 3 horizontal and 6 vertical

strips, and contains 18 sections, each identified by ordered pairs of characters, such as (*E*, *P*) or (*O*, *C*). In each ordered pair, the first entry is a letter chosen from the bounded electromagnetic system set

$M = \{E, O, I\}$.

The second entry is a letter chosen from the geometric set

$G = \{P, S, T, N, C, A\}$.

That is, the actuator/sensor set is give as

$$M \quad G = \{(E, P), (E, S), (E, T), \cdots, (I, N), (I, C), (I, A)\}.$$

In general, we have

$$M \quad G = \{(m, g) : m \quad M \text{ and } g \quad G\}.$$

However, the geometry-electromagnetic system classifier must be extended to guarantee completeness in the synthesis of motion structures and devices. It is well-known that using the basic electromagnetic features, the following basic types of nano-, micro-, and miniscale actuators and sensors can be synthesized:

1. direct current;
2. alternating current:
   induction;
   synchronous.

That is, the actuators and sensors are classified using a type classifier

$$T = \{t : t \quad T\}.$$

It was emphasized that translational, rotational, and hybrid actuators and sensors can be synthesized, and a motion classifier is

$$= \{n : n \quad \}.$$

Therefore, we have

$$M \quad G \quad T \quad = \{(m, g, t, n) : m \quad M, g \quad G, t \quad T \text{ and } n \quad \}.$$

Winding and cooling, power and size, torque-speed characteristics, excitation and bearing, as well as other actuators/sensors distinct features can be easily distinct and classified using classifiers in order to be integrated into the synthesis, structural and performance optimization. One concludes that electromechanical actuators/sensors can be mapped by a *Z*-tuple as {electromagnetic system, geometry, type, winding, excitation, cooling, etc}.

Structural Classification Problem Solver

The algorithmic concept in the structural synthesis and performance optimization starts by selecting an initial set of competing configurations and solutions (electromagnetic system, geometry, type, et cetera) for a particular problem using specifications and requirements imposed. The solutions can be generated randomly from the entire domain, however, as was emphasized earlier, available information and accessible knowledge can be readily used in order to formulate the partial domain (classifier subset). The solutions are

evaluated for their efficiency and fitness. Performance (regret) functionals can be designed to integrate weighted cost integrands (terms), and linear and nonlinear optimization (linear and nonlinear programming) allows one to find optimal solutions. The maxima or minima can be found using the gradient-based search. Alternatively, the evolutionary algorithms can be used, and the performance functional is used to compare and rank the competing solutions. The analysis and evaluation of candidate solutions are very complex problems due to infinite number of possible solutions (it is very difficult or impossible to find solutions randomly from the entire classifier domain of all possible solutions). Thus, the solutions must be examined in the partial domain (subset) of most efficient and suitable solutions. This will allow one to define the partial classifier domain of solutions generation. Hence, the following should be performed to simplify the search and optimize the algorithm to solve a wide variety of structural synthesis and optimization problems:

formulate and apply rules and criteria for solution sustaining based upon performance analysis, assessments, and outcomes;

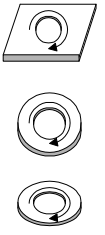
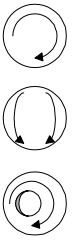
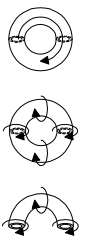
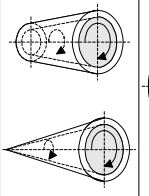
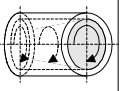
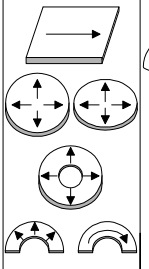
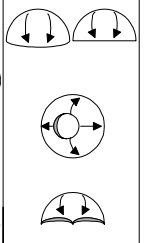
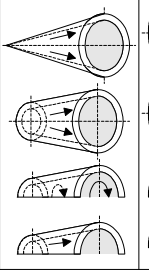
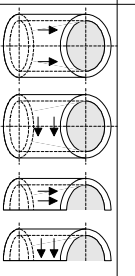
develop and generate the partial classifier domain (subset), select solution representations;

initialize solutions;

analyze and compare solutions, find the optimal solution.

Using the classifier developed, the designer can synthesize nano-, micro- and miniscale actuators and sensors (motion structures and devices which in conventional electromechanical systems terminology are called electromechanical motion devices or transducers). As an example, the structural synthesis of a two-phase permanent-magnet synchronous microscale actuator (motor) – sensor (generator) with *endless* electromagnetic system is performed using the *endless* magnetic system. The spherical, spherical-conical, conical, cylindrical, and asymmetrical geometries of the synthesized actuator/sensor are documented in Figure 3.2.1.

Table 3.2.1. Classification of nano-, micro-, and mini-scale actuators and sensors (motion structures)

M \ G	Plate, P	Spherical, S	Torroidal, T	Conical, N	Cylindrical, C	Asymmetrical,A
Electromagnetic System: *Endless* (Closed), E						○
Electromagnetic System: *Open-Ended* (Open), O						○
Electromagnetic System: *Integrated*, I	○	○	○	○	○	○

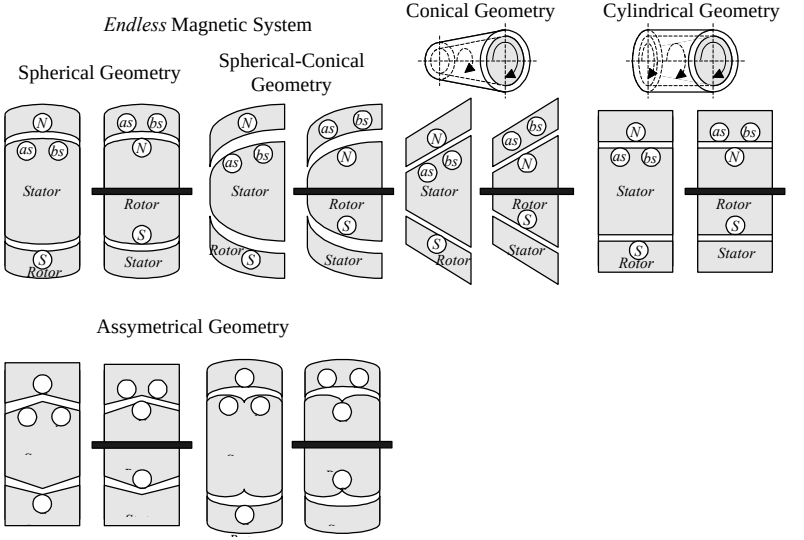


Figure 3.2.1. Synthesized microscale actuator/sensor geometry

To solve analysis, prediction, classification, modeling, and optimization problems, neural networks or genetic algorithms can be efficiently used. Neural networks and generic algorithms have evolved to the mature concepts which allow the designer to perform reliable analysis, design, and optimization. Qualitative reasoning in the structural synthesis and optimization of NEMS and MEMS is based upon artificial intelligence, and the ultimate goal is to analyze, model, and optimize qualitative models of NEMS and MEMS when knowledge, processes, and phenomena are not precisely known due to uncertainties (for example, micromachined motion microstructures properties and characteristics, e.g., charge density, thermal noise and geometry, are not precisely known and varying). It is well known that qualitative models are more reliable compared with traditional models if there is a need to perform qualitative analysis, modeling, design, optimization, and prediction. Quantitative analysis and design use a wide range of physical laws and mathematical methods to guarantee validity and robustness using partially available quantitative information.

Structural synthesis and performance optimization can be based on the knowledge domain. Qualitative representations and compositional (geometric) modeling are used to create control knowledge (existing knowledge, modeling and analysis assumptions, specific plans and requirements domains, task domain, and preferences) for solving a wide range of problems. The solving architectures are based upon qualitative reusable fundamental domains (physical laws). Qualitative reasoning must be applied to solve complex physics problems in NEMS and MEMS, as well as to perform engineering analysis and design. Emphasizing the heuristic concept for choosing the initial domain of solutions, the knowledge domain is available to efficiently and flexibly map all essential phenomena, effects, and performances. In fact, Structural Classification Table 3.2.1 ensures modeling, synthesis, and optimization in qualitative and quantitative knowledge domains carrying out numerical and analytical analysis of NEMS and MEMS. To avoid excessive computations, optimal structures can be found using qualitative analysis and design. That is, qualitative representations and compositional structural modeling can be used to create control knowledge in order to solve fundamental and engineering problems efficiently. The Structural Classification Problem Solver, which gives knowledge domain using compositional structural modeling, analysis, and synthesis can be developed applying qualitative representations. This Structural Classification Problem Solver must integrate modeling and analysis assumptions, expertise, structures, and preferences that are used in constraining search (initial structural domain). Structural optimization is given in terms of qualitative representations and compositional modeling, making fundamental concepts of the domain explicit. The Structural Classification Problem Solver can be verified solving problems analytically and numerically. Heuristic synthesis strategies and knowledge regarding physical principles must be augmented for designing nano- and

microstructures as well as nano- and microdevices. Through qualitative analysis and design, one constrains the search domain, the solutions are automatically generated, and the major performance characteristics and end-to-end behavior are predicted. Existing knowledge, specific plans and requirements domains, task domain, preferences and logical relations, make it possible to reason about the modeling and analysis assumptions explicitly, which is necessary to successfully solve fundamental and engineering problem.

The Venn diagram provides a way to represent information about NEMS and MEMS structures and configurations. Once can use regions labeled with capital letters to represent sets and use lowercase letters to represent elements. By constructing a diagram that represents some initial sets, the designer can deduce other important relations. The basic conventional form of the Venn diagram is three intersecting circles as shown in Figure 3.2.2. In this diagram, each of the circles represents a set of elements that have some common property or characteristic. Let A stands for actuators, B stands for sensors, and C stands for translational motion microstructure. Then, the region ABC represents actuators and sensors which are synthesized using translational motion microstructure, while BC is the sensor with translational motion microstructure (e.g., *i*MEMS accelerometer studied early).

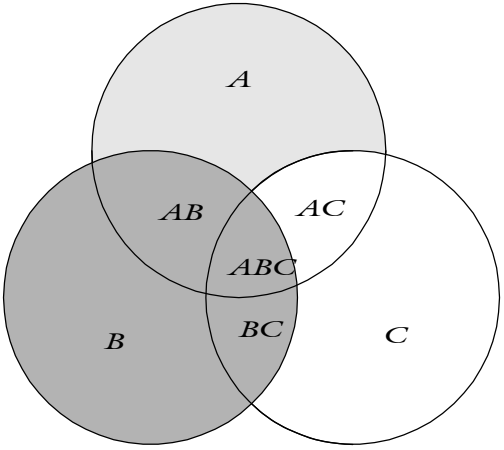


Figure 3.2.2. Venn diagram, $p = 3$: The closed curves are circles, and eight regions are labeled with the interiors that are included in each intersection. The eighth region is the outside, corresponding to the empty set

Let $A = \{a_1, a_2, ..., a_{p-1}, a_p\}$ is the collection of simple closed curves in the XY plane.

The collection A is said to be an independent family if the intersection of $b_1, b_2, ..., b_{p-1}, b_p$ is nonempty, where each b_i is either $int(a_i)$ (the interior of

a_i) or is *ext*(a_i) (the exterior of a_i). If, in addition, each such intersection is connected, then A is a p-Venn diagram, where p is the number of curves in the diagram.

3.2.2. Algebra of Sets

A set is a collection of objects (order is not significant and multiplicity is usually ignored) called the elements of the set. Symbols are used widely in the algebra of sets.

If a is an element of set A, we have $a \in A$, and if a is not an element of set A, we write $a \notin A$.

If a set A contains only the single element a, it is denoted as $\{a\}$.

The null set (set does not contain any elements) is denoted as $\varnothing$.

Two sets, A and B are equal ($A = B$) if $a \in A$ iff $a \in B$.

If $a \in A$ implies that $a \in B$, then A is a subset of B, and $A \subseteq B$.

The symbols $\subset$ and $\subseteq$ are used to describe a proper and an improper subsets. For example, if $A \subseteq B$ and $B \subseteq A$, then A is called an improper subset of B, $A = B$ (if there exists element b in B which is not in A, then A is a proper subset of B).

If the set of all elements under consideration make up the universal set U, then $A \subseteq U$.

The set A' is the complement of set A, if it is made up of all the elements of U which are not elements of A. For each set A there exists a unique set A' such that $A \cup A' = U$ and $A \cap A' = \varnothing$. Furthermore, $(A')' = A$.

Two operations on sets are union $\cup$ and intersection $\cap$. For example, an element $a \in A \cup B$ iff $a \in A$ or $a \in B$. In contrast, an element $a \in A \cap B$ iff $a \in A$ and $a \in B$.

Using $\cup$ and $\cap$ operators we have the following algebra of sets laws:

closure: there is a unique set $A \cup B$ which is a subset of U, and there is a unique set $A \cap B$ which is a subset of U;

commutative: $A \cup B = B \cup A$ and $A \cap B = B \cap A$;

associative:
$(A \cup B) \cup C = A \cup (B \cup C)$ and $(A \cap B) \cap C = A \cap (B \cap C)$;

distributive:
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,

and using the index set l, one has

$$A \cup \bigcap_l B_l = \bigcap_l (A \cup B_l) \text{ and } A \cap \bigcup_l B_l = \bigcup_l (A \cap B_l);$$

idempotent: $A \cup A = A$ and $A \cap A = A$;

identity: $A \bigcup \quad = A$ and $A \bigcap U = A$;

DeMorgan's: $(A \bigcup B)' = A' \bigcap B'$ and $(A \bigcap B)' = A' \bigcup B'$;

U and laws: $U \bigcup A = U$, $U \bigcap A = A$,

$\bigcup A = A$ and $\bigcap A = \quad$.

Additional rules and properties of the complement are:

$A \quad (A \bigcup B)$, $(A \bigcap B) \quad A$, $A \quad U$, $\quad A$

If $A \quad B$ then $A \bigcup B = B$, and if $B \quad A$ then $A \bigcap B = B$.

Sets and Lattices

A set is simply a collection of elements. For example, *a*, *b* and *c* can be grouped together as a set which is expressed as {*a*, *b*, *c*} where the curly braces are used to enclose the elements that constitute a set. In addition to the set {*a*, *b*, *c*} we define the sets {*a*, *b*} and {*d*, *e*, *g*}. Using the union operation, we have

$\{a,b,c\} \bigcup \{a,b\} = \{a,b,c\}$ and $\{a,b,c\} \bigcup \{d,e,g\} = \{a,b,c,d,e,g\}$,

while the intersection operation leads us to

$\{a,b,c\} \bigcap \{a,b\} = \{a,b\}$ and $\{a,b,c\} \bigcap \{d,e,g\} = \quad = \{\}$,

where {} is the empty (or null) set.

The subset relation can be used to partially order a set of sets. If some set *A* is a subset of a set *B*, then these sets are partially ordered with respect to each other. If a set *A* is not a subset of set *B*, and *B* is not a subset of *A*, then these sets are not ordered with respect to each other. This relation can be used to partially order a set of sets in order to classify NEMS and MEMS. Sets possess some additional structural, geometrical, as well as other properties. Additional definitions and properties can be formulated and used applying lattices.

Using a lattice, we have

$A \quad A$ (reflexive law);

if $A \quad B$ and $B \quad A$, then *A*=*B* (antisymmetric law);

if $A \quad B$ and $B \quad C$, then $A \quad C$ (transitive law);

A and *B* have a unique greatest bound, $A \bigcap B$. Furthermore, $G = A \bigcap B$, or *G* is the greatest lower bound of *A* and *B* if: $A \quad G$, $B \quad G$, and if *W* is any lower bound of *A* and *B*, then $G \quad W$;

A and *B* have a unique least upper bound, $A \bigcup B$. Furthermore, $L = A \bigcup B$, or *L* is the least upper bound of *A* and *B* if: $L \quad A$, $L \quad B$, and if *P* is any upper bound of *A* and *B*, then $P \quad L$.

A lattice is a partially ordered set where for any pair of sets (hypotheses) there is a least upper bound and greatest lower bound. Let our current hypothesis is *H*1 and the current training example is *H*2. If *H*2 is a subset of

$H1$, then no change of $H1$ is required. If $H2$ is not a subset of $H1$, then $H1$ must be changed. The minimal generalization of $H1$ is the least upper bound of $H2$ and $H1$, and the minimal specialization of $H1$ is the greatest lower bound of $H2$ and $H1$. Thus, the lattice serves as a map that allows us to locate out current hypothesis $H1$ with reference to the new information $H2$. There exists the correspondence between the algebra of propositional logic and the algebra of sets. We refer to a hypotheses as logical expressions, as rules that define a concept, or as subsets of the possible instances constructible from some set of dimensions. Furthermore, union and intersection were the important operators used to define a lattice. In addition, the propositional logic expressions can also be organized into a corresponding lattice to implement the artificial learning.

A general structure S is an ordered pair formed by a set object O and a set of binary relations R such that

$$S = (O, R) = \bigcup_{i=1}^{n} S_i ,$$

where $O = \{o_1, o_2, \ldots, o_{z-1}, o_z\}$, o_i O; $R=\{r_1, r_2, \ldots, r_{p-1}, r_p\}$, r_i R; S_i is the simple structure.

In the set object O we define the input n, output u, and internal a variables. We have $o_i = \{q_1^i, q_2^i, ..., q_{g\ 1}^i, q_g^i\}$, $q_j^i = (n_j^i, u_j^i, a_j^i)$, q_j^i O^3. Hence, the range of q, as a subset of O, is $R(q)$. Using the input-output structural function, different NEMS and MEMS can be synthesized. The documented general theory of structural optimization, which is built using the algebra of sets, allows the designer to derive relationships, flexibly adapt, fit, and optimize the nano- and microscale structures within the sets of given possible solutions.

3.3. DIRECT-CURRENT MICROMACHINES

It has been shown that the basic electromagnetic principles and fundamental physical laws are used to design motion nano- and micro-structures. Nano- and microengineering leverages from conventional theory of electromechanical motion devices, electromagnetics, integrated circuits, and quantum mechanics. The fabrication of motion microstructures is based upon CMOS (VLSI) technology, and rotational and translational transducers (actuators and sensors) were manufactured and tested. The major challenge is the difficulties to fabricate windings for microdevices (micro electric machines), reliability and ruggedness (due to bearing problems), etc. It appears that novel fabrication technologies allow one to overcome many challenges. The most efficient class of micromachines to be used as MEMS motion microdevices are induction and synchronous. These micromachines do not have collector, and the stator windings (for induction micromachines) and permanent-magnet stator (for synchronous micromachines, e.g., permanent-magnet synchronous machine and stepper motors) have been manufactured and tested. Direct-current machines are not the preferable choice. However, these micromachines will be covered first because students and engineers are familiar with these electric machines. Furthermore, even using the conventional manufacturing technology, miniscale DC motors (less than 2 mm diameter) have been massively manufactured for pagers, phones, cameras, etc.

The list of basic variables and symbols used in this chapter is given below:

i_a is the currents in the armature winding;

u_a is the applied voltages to the armature windings;

w_r and q_r are the angular velocity and angular displacement of the rotor;

E_a is the *electromotive force*;

T_e and T_L are the electromagnetic and load torques;

r_a is the resistances of the armature windings;

L_a is the self-inductances of the armature windings;

B_m is the viscous friction coefficient;

J is the equivalent moment of inertial of the rotor and attached load.

Micro- and miniscale permanent-magnet electric motors and generators are rotating energy-transfer electromechanical motion devices which convert energy by means of rotational motions. Electric machines (motion structures and motion devices) are the major part of MEMS, and therefore they must be thoroughly studied with the driving ICs. Electric motors convert electrical

energy to mechanical energy, while generators convert mechanical energy to electrical energy. It is worth mentioning that permanent-magnet electric machines can be used as motors and generators. Hence, the energy conversion is reversible, and conventional generators can be operated as motors and vice versa. That is, micro permanent-magnet electric machines can be used as the actuators and sensors. Electric machines have stationary and rotating members, separated by an air gap. The armature winding is placed in the rotor slots and connected to a rotating commutator, which rectifies the induced voltage, see Figure 3.3.1. One supplies the armature and excitation voltages or feds the armature and excitation currents to the armature (rotor) and field (stator) windings. These stator and rotor windings are coupled magnetically.

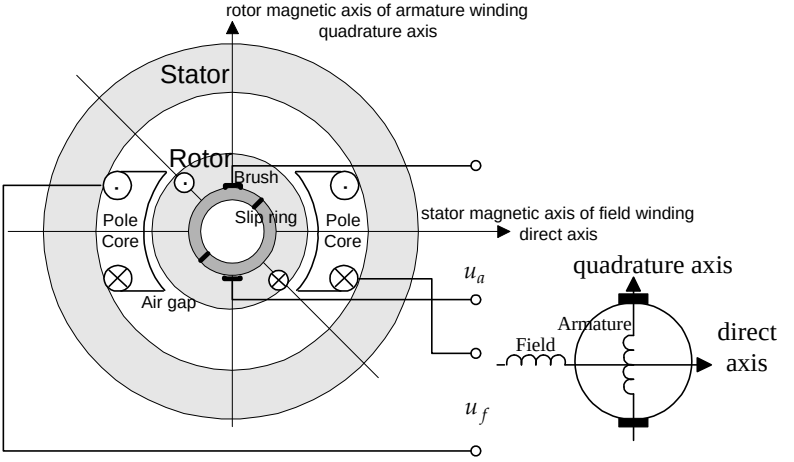


Figure 3.3.1. Two-pole DC machine with commutator

The brushes, which are connected to the armature windings, ride on the commutator. The armature winding consists of identical coils carried in slots uniformly distributed around the periphery of the rotor. The stator has salient poles, and conventional DC machines are excited by direct current; in particular, if a voltage-fed converter is used, a dc voltage u_f is supplied to the stationary field windings. Hence, the excitation magnetic field is produced by the field coils. It should be emphasized that in the permanent-magnet machines, the magnetic field is established by permanent magnets. Due to the commutator (circular conducting segments), armature and field windings produce stationary *magnetomotive forces* which are displaced by 90 electrical degrees. The armature magnetic force is along the *quadrature* (rotor) magnetic axis, while the *direct* axis stands for a field magnetic axis. The electromagnetic torque is produced as a result of the interaction of these stationary *magnetomotive forces.*

From Kirchhoff's law, one obtains the following steady-state equation for the armature voltage for electric motors (the armature current opposes the induced *electromotive force*)

$$u_a \quad E_a = i_a r_a .$$

For generators, the armature current is in the same direction as the generated *electromotive force*, and we have

$$u_a \quad E_a = \quad i_a r_a .$$

The difference between the applied voltage and the induced *electromotive force* is the voltage drop across the internal armature resistance r_a. One concludes that electric machines rotate at an angular velocity at which the *electromotive force* generated in the armature winding balances the armature voltage. If an electric machine operates as a motor, the induced *electromotive force* is less than the voltage applied to the windings. If an electric machine operates as a generator, the generated (induced) *electromotive force* is greater than the terminal voltage.

The constant magnetic flux in AC and DC machines can be produced by permanent-magnet poles. Electric machines with permanent-magnet poles are called permanent-magnet machines. Permanent-magnet DC motor/generator is illustrated in Figure 3.3.2, and a schematic diagram of permanent-magnet DC machines is illustrated in Figure 3.3.3.

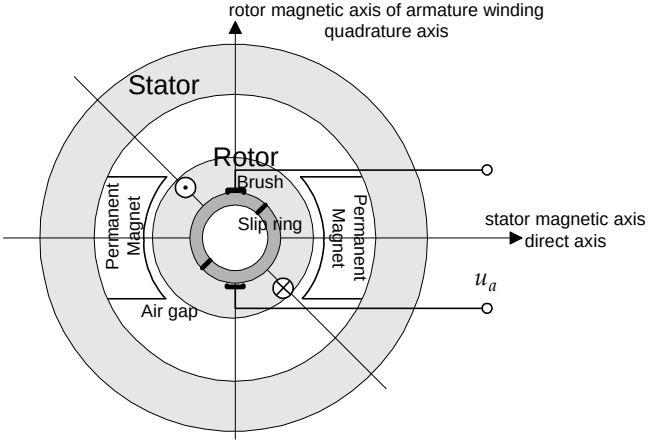


Figure 3.3.2. Permanent-magnet DC machines

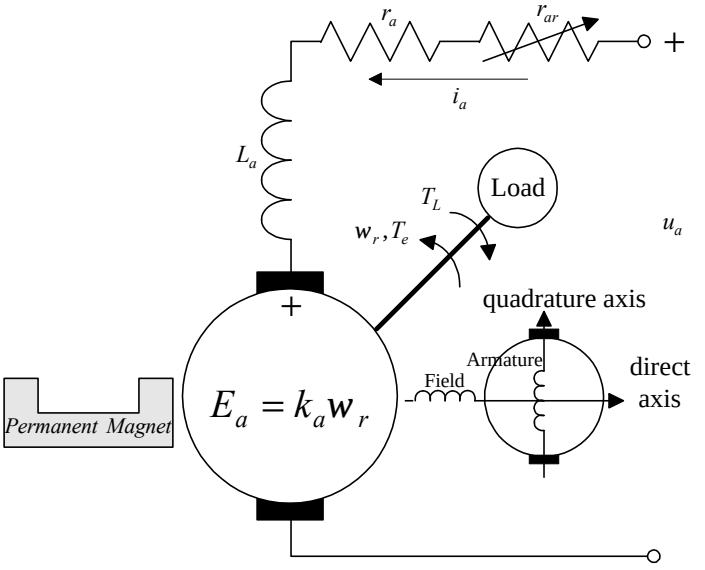


Figure 3.3.3. Schematic diagram of permanent-magnet electric machines (current direction corresponds to the motor operation)

By using Kirchhoff's voltage law and Newton's second law of motion, the differential equations for permanent-magnet DC machines can be easily derived. Assuming that the *susceptibility* is constant (in reality, Curie's constant varies as a function of temperature), one supposes that the flux, established by the permanent magnet poles, is constant. Then, denoting the *back emf* and *torque* constants as k_a, we have the following differential equations describing the transient behavior of the armature winding and *torsional-mechanical* dynamics

$$\frac{di_a}{dt} = \frac{r_a}{L_a} i_a \quad \frac{k_a}{L_a} w_r + \frac{1}{L_a} u_a ,$$

$$\frac{dw_r}{dt} = \frac{k_a}{J} i_a \quad \frac{B_m}{J} w_r \quad \frac{1}{J} T_L . \qquad (3.3.1)$$

From (3.3.1) it is easy to find the model in matrix form. We have

$$\begin{matrix} \frac{di_a}{dt} \\ \frac{dw_r}{dt} \end{matrix} = \begin{matrix} \frac{r_a}{L_a} & \frac{k_a}{L_a} \\ \frac{k_a}{J} & \frac{B_m}{J} \end{matrix} \begin{matrix} i_a \\ w_r \end{matrix} + \begin{matrix} \frac{1}{L_a} & 0 \\ 0 & \frac{1}{J} \end{matrix} \begin{matrix} u_a \\ T_L . \end{matrix}$$

An *s*-domain block diagram of permanent-magnet DC motors is illustrated in Figure 3.3.4.

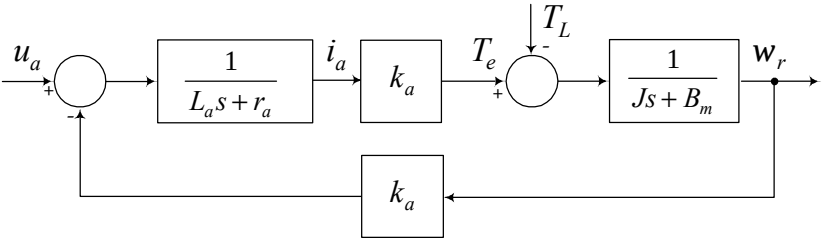


Figure 3.3.4. Block diagram of permanent-magnet DC motors: *s*-domain block diagram

The angular velocity can be reversed if the polarity of the applied voltage is changed (the direction of the field flux cannot be changed). The steady-state torque-speed characteristic curves obey the following equation

$$w_r = \frac{u_a}{k_a} - \frac{r_a i_a}{k_a} = \frac{u_a}{k_a} - \frac{r_a}{k_a^2} T_e, \tag{3.3.2}$$

and a spectrum of the torque-speed characteristic curves is illustrated in Figure 3.3.5.

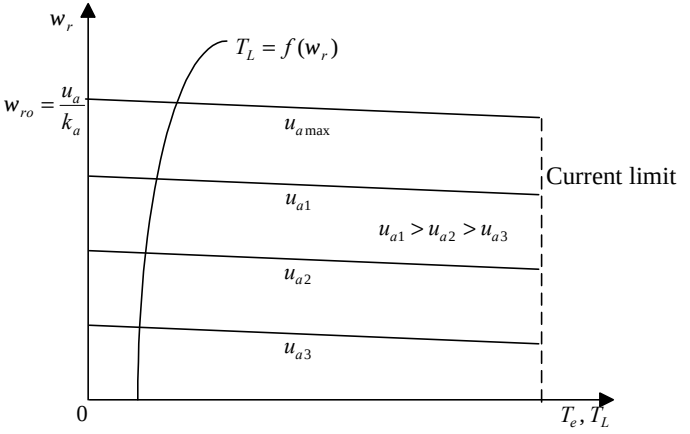


Figure 3.3.5.Torque-speed characteristics for permanent-magnet motors

If micro permanent-magnet DC machine is used as the generator (tachogenerator measure the angular velocity), the circuitry dynamics for the resistive load R_L is given as

$$\frac{di_a}{dt} = -\frac{r_a + R_L}{L_a} i_a + \frac{k_a}{L_a} w_r.$$

That is, in the steady-state, the armature current is proportional to the angular velocity, and we have

$$i_a = \frac{k_a}{r_a + R_L} w_r .$$

As was emphasized in section 1.5, flip-chip MEMS are found wide application due to low cost and application of well-developed fabrication processes. For example, monolithic dual power operational amplifiers (Motorola TCA0372 available in plastic packages 751G, 648, and 626) feeds DC minimotor to regulate the angular velocity, see Figure 3.3.6.

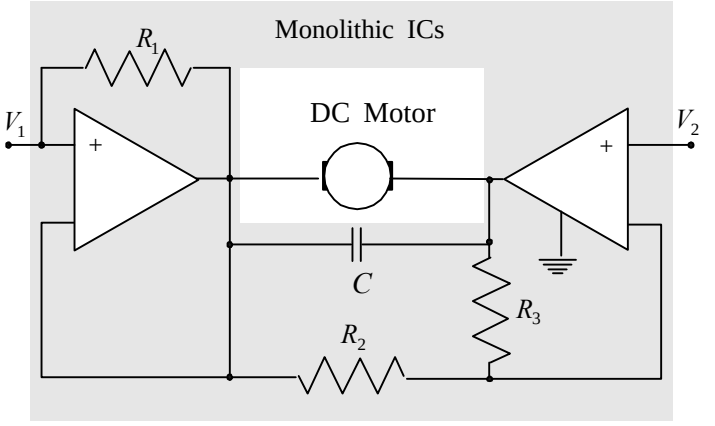


Figure 3.3.6. Application of a monolithic IC to control DC motor

Motion micro- and ministructures as well as micro- and minidevices (actuators and sensors) are mounted face down with bumps on the pads that form electrical and mechanical joints to the ICs substrate. Figure 3.3.7 illustrates a flip-chip MEMS with permanent-magnet micromotor driven by the MC33030 monolithic servo motor driver. Control algorithms are implemented to control the angular velocity of electric motors. The MC33030 integrates on-chip operational amplifier and comparator, driving and braking logic, PWM four-quadrant converter, etc. The MC33030 data and complete description are given. As in the conventional configurations, the difference between the reference (command) and actual angular velocity or displacement, linear velocity or position, is compared by the *error amplifier*, and two comparators are used. A *pnp* differential output power stage provides driving and braking capabilities, and the four-quadrant H-configured power stage guarantees high performance and efficiency. Using the error between the desired (command) and actual angular velocity or displacement, the bipolar voltage u_a is applied to the armature winding, the electromagnetic torque is developed due to the current i_a, and micro- or minimotor rotates. A complete description of the MC33030 monolithic servomotors driver is given below.

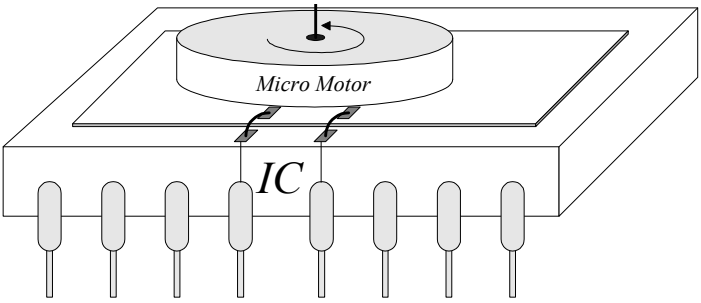


Figure 3.3.7. Flip-chip monolithic MEMS: MC33030 ICs and micromotor

(Copyright of Motorola, used with permission)

MOTOROLA

MC33030

DC Servo Motor Controller/Driver

The MC33030 is a monolithic DC servo motor controller providing all active functions necessary for a complete closed loop system. This device consists of an on–chip op amp and window comparator with wide input common–mode range, drive apd brake logic with direction memory, Power H–Switch driver capable of 1.0 A, independently programmable over–current monitor and shutdown delay, and over–voltage monitor. This part is ideally suited for almost any servo positioning application that requires sensing of temperature, pressure, light, magnetic flux, or any other means that can be converted to a voltage.

Although this device is primarily intended for servo applications, it can be used as a switchmode motor controller.

- On–Chip Error Amp for Feedback Monitoring
- Window Detector with Deadband and Self Centering Reference Input
- Drive/Brake Logic with Direction Memory
- 1.0 A Power H–Switch
- Programmable Over–Current Detector
- Programmable Over–Current Shutdown Delay
- Over–Voltage Shutdown

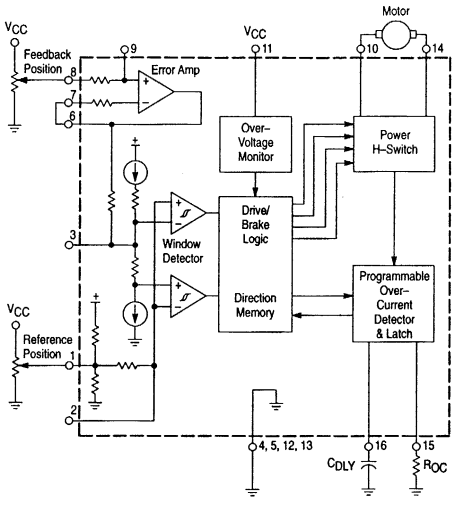
DC SERVO MOTOR CONTROLLER/DRIVER

SEMICONDUCTOR TECHNICAL DATA

P SUFFIX
PLASTIC PACKAGE
CASE 648C
(DIP–16)

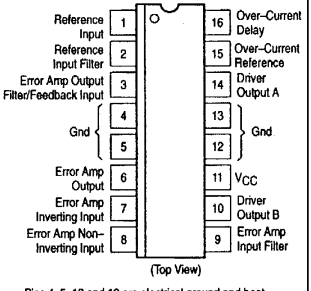
DW SUFFIX
PLASTIC PACKAGE
CASE 751G
(SOP–16L)

Representative Block Diagram



This device contains 119 active transistors.

PIN CONNECTIONS



(Top View)

Pins 4, 5, 12 and 13 are electrical ground and heat sink pins for IC.

ORDERING INFORMATION

Device	Operating Temperature Range	Package
MC33030DW	$T_A = -40°$ to $+85°C$	SOP–16L
MC33030P		DIP–16

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Power Supply Voltage	V_{CC}	36	V
Input Voltage Range Op Amp, Comparator, Current Limit (Pins 1, 2, 3, 6, 7, 8, 9, 15)	V_{IR}	−0.3 to V_{CC}	V
Input Differential Voltage Range Op Amp, Comparator (Pins 1, 2, 3, 6, 7, 8, 9)	V_{IDR}	−0.3 to V_{CC}	V
Delay Pin Sink Current (Pin 16)	$I_{DLY\ (sink)}$	20	mA
Output Source Current (Op Amp)	I_{source}	10	mA
Drive Output Voltage Range (Note 1)	V_{DRV}	−0.3 to (V_{CC} + V_F)	V
Drive Output Source Current (Note 2)	$I_{DRV(source)}$	1.0	A
Drive Output Sink Current (Note 2)	$I_{DRV(sink)}$	1.0	A
Brake Diode Forward Current (Note 2)	I_F	1.0	A
Power Dissipation and Thermal characteristics P Suffix, Dual In Line Case 648C Thermal Ressistance, Junction–to–Air Thermal Ressistance, Junction–to–Case (Pins 4, 5, 12, 13) DW Suffix, Dual In Line Case 751G Thermal Ressistance, Junction–to–Air Thermal Ressistance, Junction–to–Case (Pins 4, 5, 12, 13)	 $R_{\theta JA}$ $R_{\theta JC}$ $R_{\theta JA}$ $R_{\theta JC}$	 80 15 94 18	°C/W
Operating Junction Temperature	T_J	+150	°C
Operating Ambient Temperature Range	T_A	−40 to +85	°C
Storage Temperature Range	T_{stg}	−65 to +150	°C

NOTES: 1. The upper voltage level is clamped by the forward drop, V_F of the brake diode.
2. These values are for continuous DC current. Maximum package power dissipation limits must be observed.

ELECTRICAL CHARACTERISTICS (V_{CC} = 14 V, T_A = 25°C, unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
ERROR AMP					
Input Offset Voltage (−40°C ≤ T_A ≤ 85°C) $V_{Pin\ 6}$ = 7.0 V, R_L = 100 k	V_{IO}	–	1.5	10	mV
Input Offset Current ($V_{Pin\ 6}$ = 1.0 V, R_L = 100 k)	I_{IO}	–	0.7	–	nA
Input Bias current ($V_{Pin\ 6}$ = 1.0 V, R_L = 100 k)	I_{IB}	–	7.0	–	nA
Input Common–Mode Voltage Range ΔV_{IO} = 20 mV, R_L = 100 k	V_{ICR}	–	0 to (V_{CC} − 1.2)	–	V
Slew Rate, Open Loop (V_{ID} = 0.5 V, C_L = 15 pF)	SR	–	0.40	–	V/μs
Unity–Gain Crossover Frequency	f_C	–	550	–	kHz
Unity–Gain Phase Margin	ϕm	–	63	–	deg.
Common–mode Rejection Ratio ($V_{Pin\ 6}$ = 7.0 V, R_L = 100 k)	CMRR	50	82	–	dB
Power Supply Rejection Ratio V_{CC} = 9.0 to 16 V, $V_{Pin\ 6}$ = 7.0 V, RL = 100 k	PSRR	–	89	–	dB
Output Source Current ($V_{Pin\ 6}$ = 12 V)	I_{O+}	–	1.8	–	mA
Output Sink Current ($V_{Pin\ 6}$ = 1.0 V)	I_{O-}		250	–	μA
Output Voltage Swing (R_L = 17 k to Ground)	V_{OH} V_{OL}	12.5 –	13.1 0.02	– –	V V

NOTES: 3. The upper or lower hysteresis will be lost when operating the Input, Pin 3, close to the respective rail. Refer to Figure 4.
4. Low duty cycle pulse techniques are used during test to maintain junction temperature as close to ambient temperature as possible.

ELECTRICAL CHARACTERISTICS (continued) (V_{CC} = 14 V, T_A = 25°C, unless otherwise noted.)

Characteristic	Symbol	Min	Typ	Max	Unit
WINDOW DETECTOR					
Input Hysteresis Voltage ($V_1 - V_4$, $V_2 - V$, Figure 18)	V_H	25	35	45	mV
Input Dead Zone Range ($V_2 - V_4$, Figure 18)	V_{IDZ}	166	210	254	mV
Input Offset Voltage (|[$V_2 - V_{Pin\ 2}$] – [$V_{Pin\ 2} - V_4$]| Figure 18)	V_{IO}	–	25	–	mV
Input Functional Common–Mode Range (Note 3) Upper Threshold Lower Threshold	 V_{IH} V_{IL}	 – –	 (V_{CC}–1.05) 0.24	 – –	V
Reference Input Self Centering Voltage Pins 1 and 2 Open	V_{RSC}	–	(1/2 V_{CC})	–	V
Window Detector Propagation Delay Comparator Input, Pin 3, to Drive Outputs V_{ID} = 0.5 V, $R_{L(DRV)}$ = 390 Ω	t_p(IN/DRV)	–	2.0	–	μs
OVER–CURRENT MONITOR					
Over–Current References Resistor Voltage (pin 15)	R_{OC}	3.9	4.3	4.7	V
Delay Pin Source Current	$I_{DLY(SOURCE)}$	–	5.5	6.9	μA
Delay Pin Sink Current (R_{OC} = 27 k, I_{DRV} = 0 mA) V_{DLY} = 5.0 V V_{DLY} = 8.3 V V_{DLY} = 14 V	I_{DLY} (sink)	 – – –	 0.1 0.7 16.5	 – – –	mA
Delay Pin voltage, Low State (I_{DLY} = 0 mA)	$V_{OL(DLY)}$	–	0.3	0.4	V
Over–Current Shutdown Threshold V_{CC} = 14 V V_{CC} = 8.0 V	$V_{th(OC)}$	 6.8 5.5	 7.5 6.0	 8.2 6.5	V
Over–Current Shutdown Propagation Delay Delay capacitor Input, Pin 16, to Outputs, V_{ID} = 0.5 V	$t_{p(DLY/DRV)}$	–	1.8	–	μs
POWER H–SWITCH					
Drive–Output Saturation (–40°C ≤ TA ≤ +85°C, Note 4) High–State (I_{source} = 100 mA) Low–State (I_{sink} = 100 mA)	 $V_{OH(DRV)}$ $V_{OL(DRV)}$	 (V_{CC} – 2) –	 (V_{CC} – 0.85) 0.12	 – 1.0	V
Drive–Output Voltage Switiching Time (C_L = 15 pF) Rise Time Fall Time	 t_r t_f	 – –	 200 200	 – –	ns
Brake Diode Forward Voltage Drop (I_F = 200 mA, Note 4)	V_F	–	1.04	2.5	V
TOTAL DEVICE					
Standby Supply Current	I_{CC}	–	14	25	mA
Over–Voltage Shutdown Threshold (–40°C ≤ T_A ≤ +85°C)	$V_{th(OV)}$	16.5	18	20.5	V
Over–Voltage Shutdown Hysteresis (Device “off” to “on”)	$V_{H(OV)}$	0.3	0.6	1.0	V
Operating Voltage Lower Threshold (–40°C ≤ T_A ≤ +85°C)	V_{CC}	–	7.5	8.0	V

NOTES: 3. The upper or lower hysteresis will be lost when operating the Input, Pin 3, close to the respective rail. Refer to Figure 4.
4. Low duty cycle pulse techniques are used during test to maintain junction temperature as close to ambient temperature as possible.

Figure 1. Error Amp Input Common–Mode Voltage Range versus Temperature

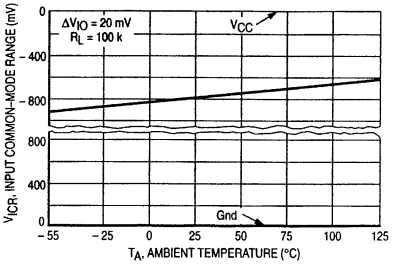


Figure 2. Error Amp Output Saturation versus Load Current

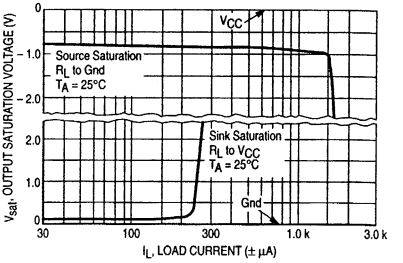


Figure 3. Open Loop Voltage Gain and Phase versus Frequency

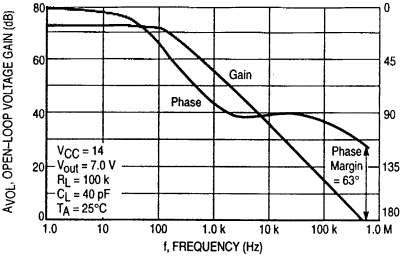


Figure 4. Window Detector Reference–Input Common–Mode Voltage Range versus Temperature

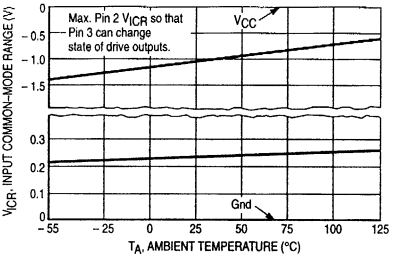

Max. Pin 2 VICR so that Pin 3 can change state of drive outputs.
VCC
Gnd
VICR, INPUT COMMON–MODE RANGE (V)
TA, AMBIENT TEMPERATURE (°C)

Figure 5. Window Detector Feedback–Input Thresholds versus Temperature

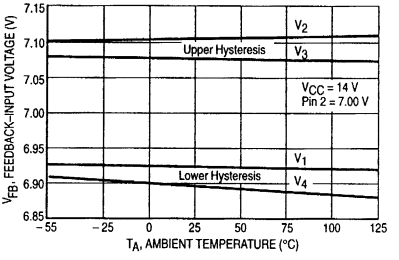


Figure 6. Output Driver Saturation versus Load Current

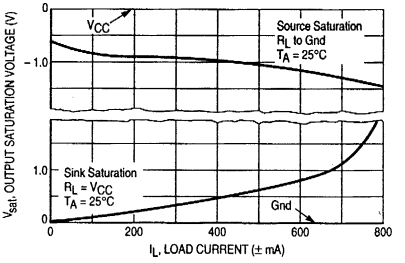


Figure 7. Brake Diode Forward Current versus Forward Voltage

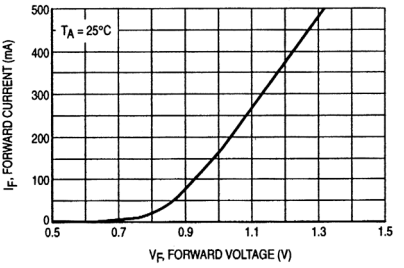


Figure 8. Output Source Current–Limit versus Over–Current Reference Resistance

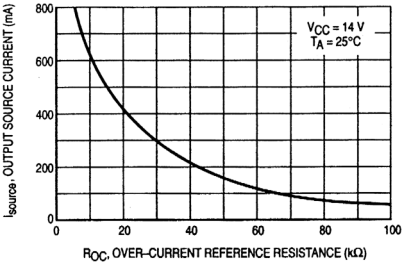


Figure 9. Output Source Current–Limit versus Temperature

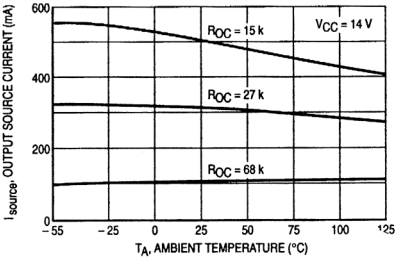


Figure 10. Normalized Delay Pin Source Current versus Temperature

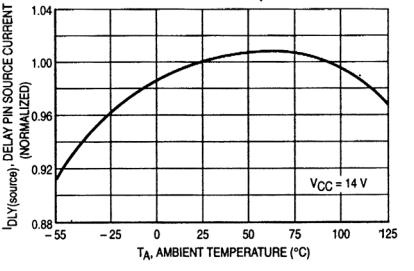


Figure 11. Normalized Over–Current Delay Threshold Voltage versus Temperature

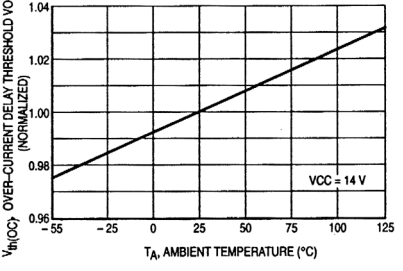


Figure 12. Supply Current versus Supply Voltage

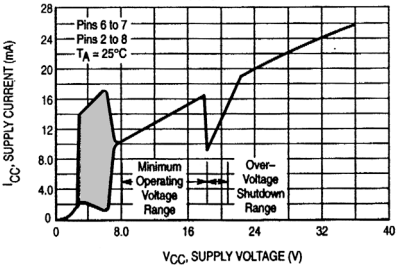


Figure 13. Normalized Over–Voltage Shutdown Threshold versus Temperature

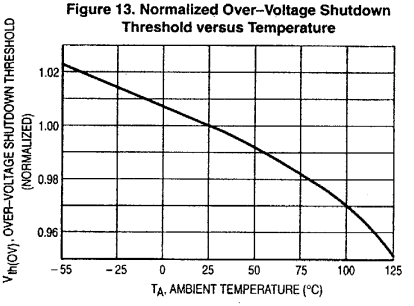


Figure 14. Normalized Over–Voltage Shutdown Hysteresis versus Temperature

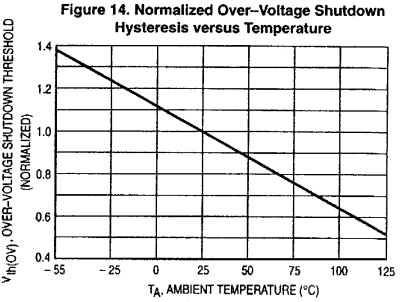


Figure 15. P Suffix (DIP–16) Thermal Resistance and Maximum Power Dissipation versus P.C.B. Copper Length

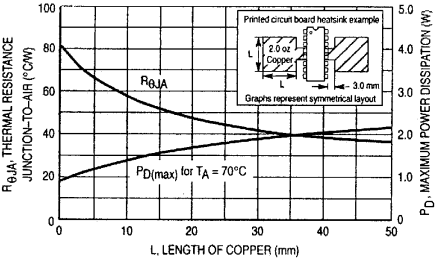
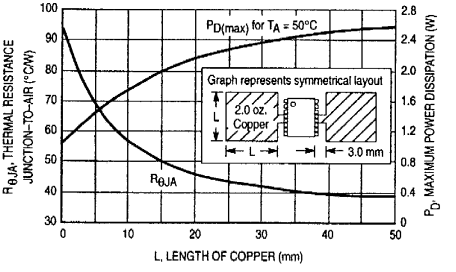


Figure 16. DW Suffix (SOP–16L) Thermal Resistance and Maximum Power Dissipation versus P.C.B. Copper Length



OPERATING DESCRIPTION

The MC33030 was designed to drive tractional horsepower DC motors and sense actuator position by voltage feedback. A typical servo application and representative internal block diagram are shown in Figure 17. The system operates by setting a voltage on the reference input of the Window Detector (Pin 1) which appears on (Pin 2). A DC motor then drives a position sensor, usually a potentiometer driven by a gear box, in a corrective fashion so that a voltage proportional to position is present at Pin 3. The servo motor will continue to run until the voltage at Pin 3 falls within the dead zone, which is centered about the reference voltage.

The Window Detector is composed of two comparators, A and B, each containing hysteresis. The reference input, common to both comparators, is pre–biased at 1/2 V_{CC} for simple two position servo systems and can easily be overriden by an external voltage divider. The feedback voltage present at Pin 3 is connected to the center of two resistors that are driven by an equal magnitude current source and sink. This generates an offset voltage at the input of each comparator which is centered about Pin 3 that can float virtually from V_{CC} to ground. The sum of the upper and lower offset voltages is defined as the window detector input dead zone range.

To increase system flexibility, an on–chip Error Amp is provided. It can be used to buffer and/or gain–up the actuator position voltage which has the effect of narrowing the dead zone range. A PNP differential input stage is provided so that the input common–mode voltage range will include ground. The main design goal of the error amp output stage was to be able to drive the window detector input. It typically can source 1.8 mA and sink 250 μA. Special design considerations must be made if it is to be used for other applications.

The Power H–Switch provides a direct means for motor drive and braking with a maximum source, sink, and brake current of 1.0 A continuous. Maximum package power dissipation limits must be observed. Refer to Figure 15 for thermal information. For greater drive current requirements, a method for buffering that maintains all the system features is shown in Figure 30.

The Over–Current Monitor is designed to distinguish between motor start–up or locked rotor conditions that can occur when the actuator has reached its travel limit. A fraction of the Power H–Switch source current is internally fed into one of the two inverting inputs of the current comparator, while the non–inverting input is driven by a programmable current reference. This reference level is controlled by the resistance value selected for R_{OC}, and must be greater than the required motor run–current with its mechanical load over temperature; refer to Figure 8. During an over–current condition, the comparator will turn off and allow the current source to charge the delay capacitor, C_{DLY}. When C_{DLY} charges to a level of 7.5 V, the set input of the over–current latch will go high, disabling the drive and brake functions of the Power H–Switch. The programmable time delay is determined by the capacitance value–selected for C_{DLY}.

$$t_{DLY} = \frac{v_{ref} C_{DLY}}{I_{DLY\,(source)}} = \frac{7.5\, C_{DLY}}{5.5\,\mu A} = 1.36\ C_{DLY}\ \text{in}\ \mu F$$

This system allows the Power H–Switch to supply motor start–up current for a predetermined amount of time. If the rotor is locked, the system will time–out and shut–down. This feature eliminates the need for servo end–of–travel or limit switches. Care must be taken so as not to select too large of a capacitance value for C_{DLY}. An over–current condition for an excessively long time–out period can cause the integrated circuit to overheat and eventually fail. Again, the maximum package power dissipation limits must be observed. The over–current latch is reset upon power–up or by readjusting $V_{Pin\ 2}$ as to cause $V_{Pin\ 3}$ to enter or pass through the dead zone. This can be achieved by requesting the motor to reverse direction.

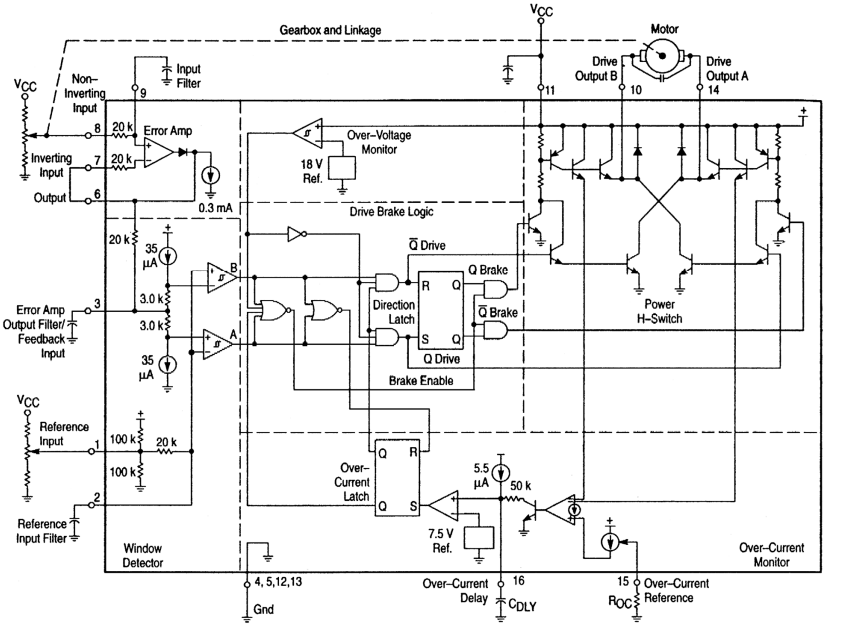
An Over–Voltage Monitor circuit provides protection for the integrated circuit and motor by disabling the Power H–Switch functions if V_{CC} should exceed 18 V. Resumption of normal operation will commence when V_{CC} falls below 17.4 V.

A timing diagram that depicts the operation of the Drive/Brake Logic section is shown in Figure 18. The waveforms grouped in [1] show a reference voltage that was preset, appearing on Pin 2, which corresponds to the desired actuator position. The true actuator position is represented by the voltage on Pin 3. The points V_1 through V_4 represent the input voltage thresholds of comparators A and B that cause a change in their respective output state. They are defined as follows:

V_1 = Comparator B turn–off threshold
V_2 = Comparator A turn–on threshold
V_3 = Comparator A turn–off threshold
V_4 = Comparator B turn–on threshold
V_1–V_4 = Comparator B input hysteresis voltage
V_2–V_3 = Comparator A input hysteresis voltage
V_2–V_4 = Window detector input dead zone range
$|(V_2–V_{Pin2}) – (V_{Pin2} – V_4)|$ = Window detector input voltage

It must be remembered that points V_1 through V_4 always try to follow and center about the reference voltage setting if it is within the input common–mode voltage range of Pin 3; Figures 4 and 5. Initially consider that the feedback input voltage level is somewhere on the dashed line between V_2 and V_4 in [1]. This is within the dead zone range as defined above and the motor will be off. Now if the reference voltage is raised so that $V_{Pin\ 3}$ is less than V_4, comparator B will turn–on [3] enabling $\overline{Q}$ Drive, causing Drive Output A to sink and B to source motor current [8]. The actuator will move in Direction B until $V_{Pin\ 3}$ becomes greater than V_1. Comparator B will turn–off, activating the brake enable [4] and $\overline{Q}$ Brake [6] causing Drive Output A to go high and B to go into a high impedance state. The inertia of the mechanical system will drive the motor as a generator creating a positive voltage on Pin 10 with respect to Pin 14. The servo system can be stopped quickly, so as not to over–shoot through the dead zone range, by braking. This is accomplished by shorting the motor/generator terminals together. Brake current will flow into the diode at Drive Output B, through the internal V_{CC} rail, and out the emitter of the sourcing transistor at Drive Output A. The end of the solid line and beginning of the dashed for $V_{Pin\ 3}$ [1] indicates the possible resting position of the actuator after braking.

Figure 17. Representative Block Diagram and Typical Servo Application



If $V_{Pin\ 3}$ should continue to rise and become greater than V_2, the actuator will have over shot the dead zone range and cause the motor to run in Direction A until $V_{Pin\ 3}$ is equal to V_3. The Drive/Brake behavior for Direction A is Identical to that of B. Overshooting the dead zone range in both directions can cause the servo system to continuously hunt or oscillate. Notice that the last motor run–direction is stored in the direction latch. This information is needed to determine whether Q or $\overline{Q}$ Brake is to be enabled when $V_{Pin\ 3}$ enters the dead zone range. The dashed lines in [8,9] indicate the resulting waveforms of an over–current condition that has exceeded the programmed time delay. Notice that both Drive Outputs go into a high impedance state until $V_{Pin\ 2}$ is readjusted so that $V_{Pin\ 3}$ enters or crosses through the dead [7,4].

The inputs of the Error Amp and Window Detector can be susceptible to the noise created by the brushes of the DC motor and cause the servo to hunt. Therefore, each of these inputs are provided with an internal series resistor and are pinned out for an external bypass capacitor. It has been found that placing a capacitor with *short leads* directly across the brushes will significantly reduce noise problems. Good quality RF bypass capacitors in the range of 0.001 to 0.1 µF may be required. Many of the more economical motors will generate significant levels of RF energy over a spectrum that extends from DC to beyond 200 MHz. The capacitance value and method of noise filtering must be determined on a system by system basis.

Thus far, the operating description has been limited to servo systems in which the motor mechanically drives a potentiometer for position sensing. Figures 19, 20, 27, and 31 show examples that use light, magnetic flux, temperature, and pressure as a means to drive the feedback element. Figures 21, 22 and 23 are examples of two position, open loop servo systems. In these systems, the motor runs the actuator to each end of its travel limit where the Over–Current Monitor detects a locked rotor condition and shuts down the drive. Figures 32 and 33 show two possible methods of using the MC33030 as a switching motor controller. In each example a fixed reference voltage is applied to Pin 2. This causes $V_{Pin\ 3}$ to be less than V_4 and Drive Output A, Pin 14, to be in a low state saturating the TIP42 transistor. In Figure 32, the motor drives a tachometer that generates an ac voltage proportional to RPM. This voltage is rectified, filtered, divided down by the speed set potentiometer, and applied to Pin 8. The motor will accelerate until $V_{Pin\ 3}$ is equal to V_1 at which time Pin 14 will go to a high state and terminate the motor drive. The motor will now coast until $V_{Pin\ 3}$ is less than V_4 where upon drive is then reapplied. The system operation of Figure 31 is identical to that of 32 except the signal at Pin 3 is an amplified average of the motors drive and back EMF voltages. Both systems exhibit excellent control of RPM with variations of V_{CC}; however, Figure 32 has somewhat better torque characteristics at low RPM.

Figure 18. Timing Diagram

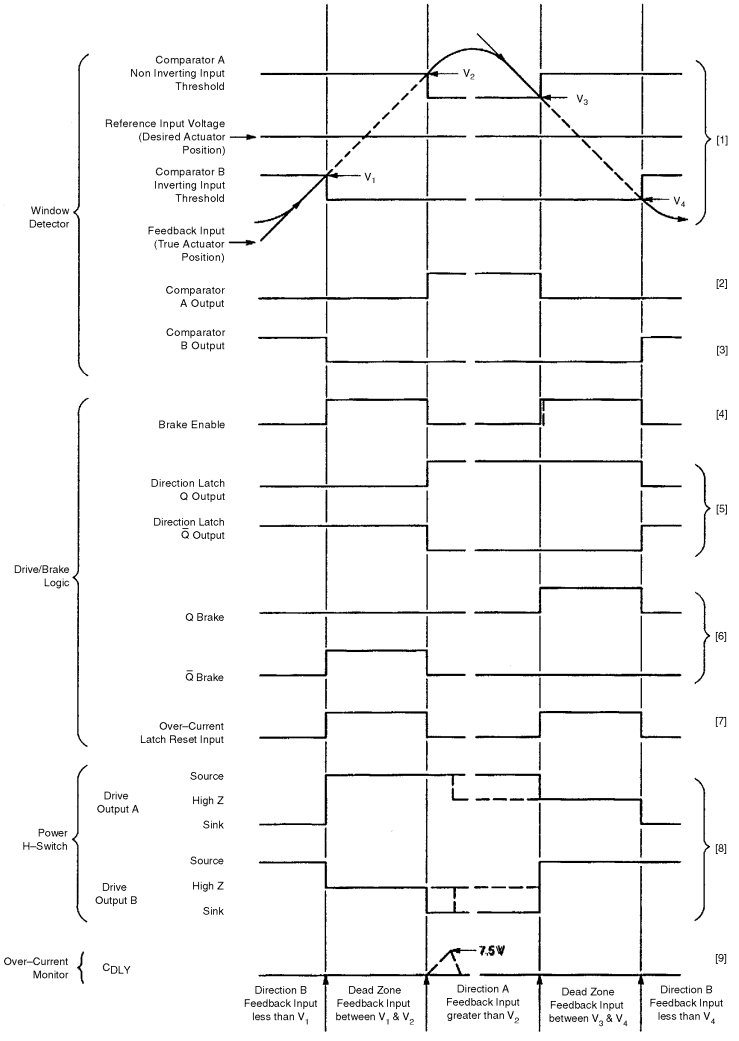


Figure 19. Solar Tracking Servo System

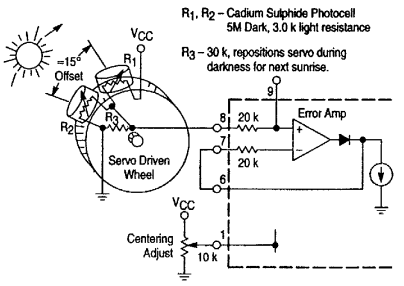
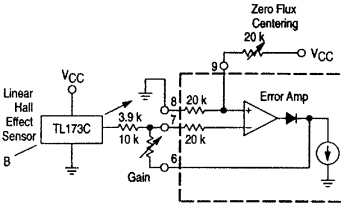
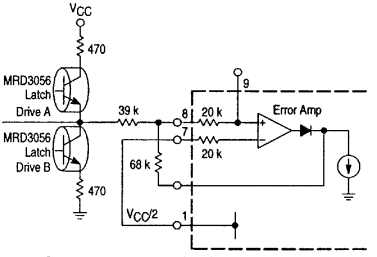


Figure 20. Magnetic Sensing Servo System



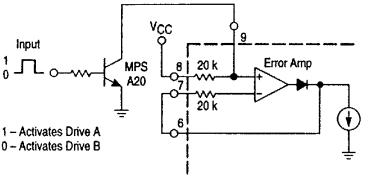
Typical sensitivity with gain set at 3.9 k is 1.5 mV/gauss.
Servo motor controls magnetic field about sensor.

Figure 21. Infrared Latched Two Position Servo System



Over–current monitor (not shown) shuts down
servo when end stop is reached.

Figure 22. Digital Two Position Servo System



Over–current monitor (not shown) shuts down
servo when end stop is reached.

Figure 23. 0.25 Hz Square–Wave Servo Agitator

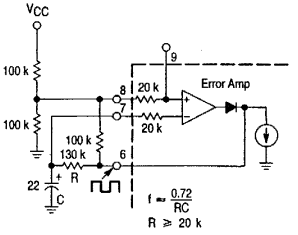


Figure 24. Second Order Low–Pass Active Filter

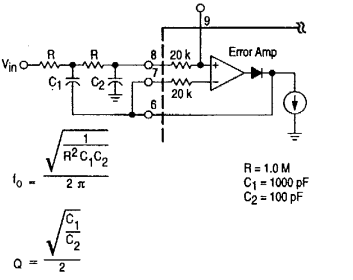
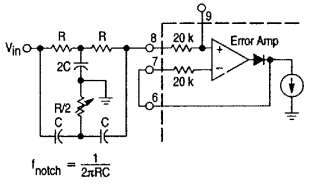


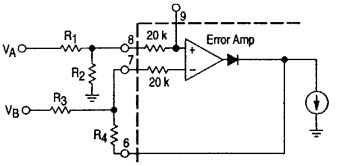
Figure 25. Notch Filter



$$f_{notch} = \frac{1}{2\pi RC}$$

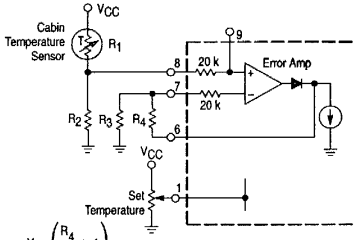
For 60 Hz R = 53.6 k, C = 0.05

Figure 26. Differential Input Amplifier



$$V_{Pin\ 6} = V_A\left(\frac{R_3 + R_4}{R_1 + R_2}\right)\frac{R_2}{R_3} - \left(\frac{R_4}{R_3}\,V_B\right)$$

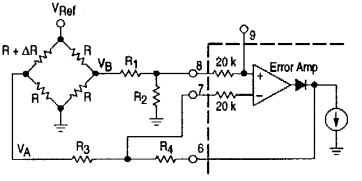
Figure 27. Temperature Sensing Servo System



$$V_{Pin\ 6} = \frac{V_{CC}\left(\frac{R_4}{R_3} + 1\right)}{\left(\frac{R_1}{R_2} + 1\right)}$$

In this application the servo motor drives the heat/air conditioner modulator door in a duct system.

Figure 28. Bridge Amplifier

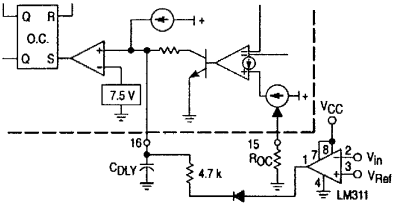


$$V_A - V_B = V_{Ref}\left(\frac{\Delta R}{4R + 2\Delta R}\right)$$

$R_1 = R_3$, $R_2 = R_4$, $R_1 >> R$

$$V_{Pin\ 6} = \frac{R_4}{R_3}(V_A - V_B)$$

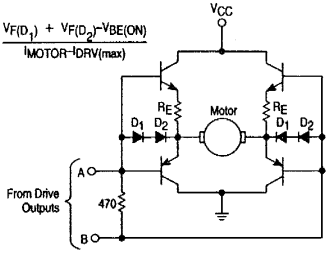
Figure 29. Remote Latched Shutdown



A direction change signal is required at Pins 2 or 3 to reset the over−current latch.

Figure 30. Power H−Switch Buffer

$$R_E \cong \frac{V_{F(D_1)} + V_{F(D_2)} - V_{BE(ON)}}{I_{MOTOR} - I_{DRV(max)}}$$



This circuit maintains the brake and over−current features of the MC33030. Set R_{OC} to 15 k for $I_{DRV(max)} \approx 0.5$ A.

Figure 31. Adjustable Pressure Differential Regulator

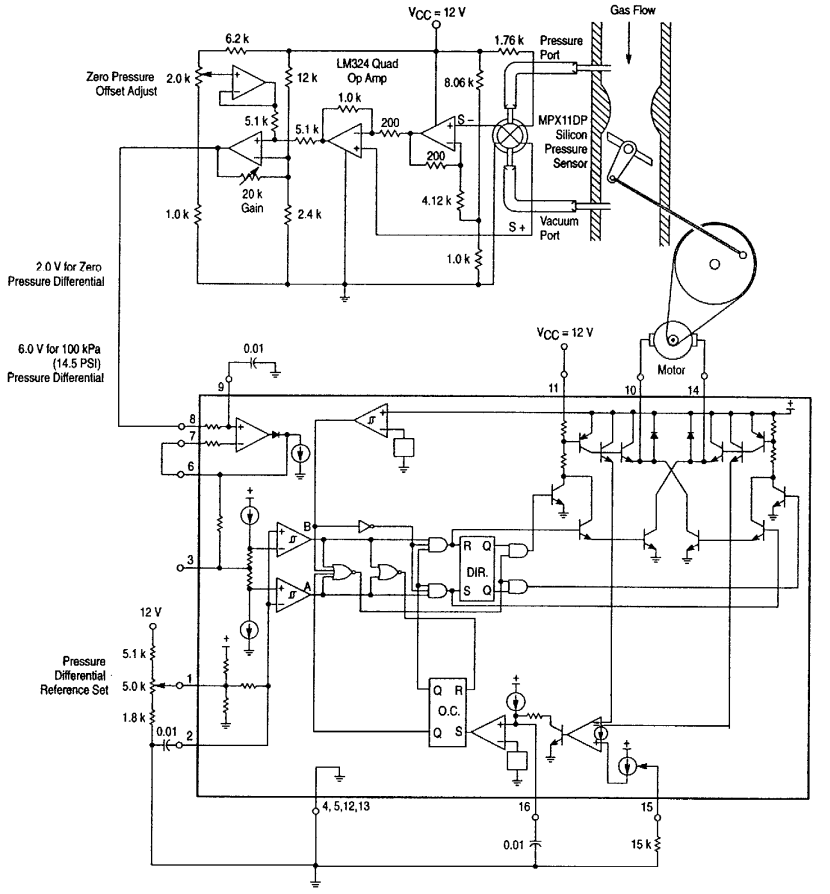


Figure 32. Switching Motor Controller With Buffered Output and Tach Feedback

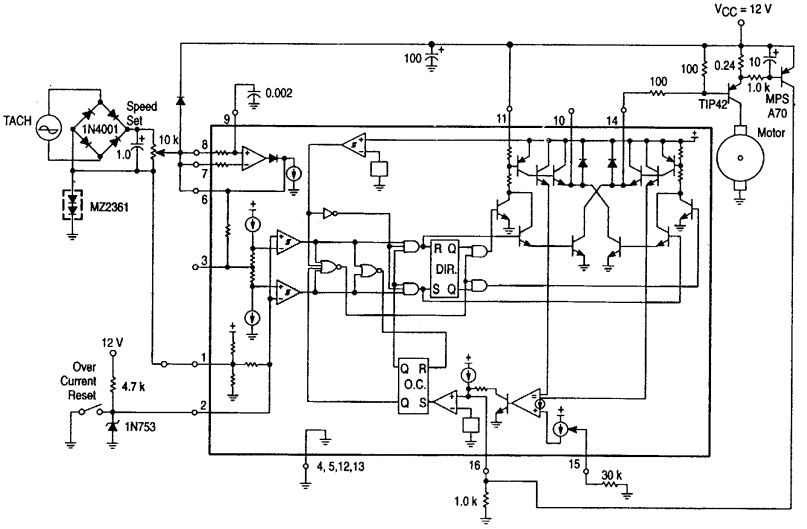
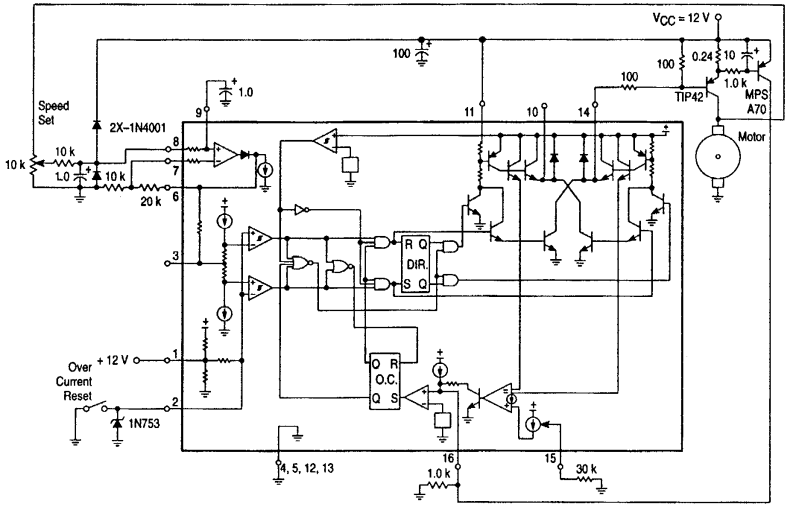


Figure 33. Switching Motor Controller With Buffered Output and Back EMF Sensing



CHAPTER 4

CONTROL OF NANO- AND MICROELECTROMECHANICAL SYSTEMS

4.1. FUNDAMENTALS OF ELECTROMAGNETIC RADIATION AND ANTENNAS IN NANO- AND MICROSCALE ELECTROMECHANICAL SYSTEMS

The electromagnetic power is generated and radiated by antennas. Time-varying current radiates electromagnetic waves (radiated electromagnetic fields). Radiation pattern, beam width, directivity, and other major characteristics can be studied using Maxwell's equations, see Section 2.2. We use the vectors of the electric field intensity $\boldsymbol{E}$, electric flux density $\boldsymbol{D}$, magnetic field intensity $\boldsymbol{H}$, and magnetic flux density $\boldsymbol{B}$. The constitutive equations are

$$\mathbf{D} = e\mathbf{E} \text{ and } \mathbf{B} = m\mathbf{H}$$

where e is the permittivity; m is the permiability.

It was shown in Section 2.2 that in the static (time-invariant) fields, electric and magnetic field vectors form separate and independent pairs. That is, $\boldsymbol{E}$ and $\boldsymbol{D}$ are not related to $\boldsymbol{H}$ and $\boldsymbol{B}$, and vice versa. However, for time-varying electric and magnetic fields, we have the following fundamental electromagnetic equations

$$\mathbf{E}(x,y,z,t) = -m\frac{\partial \mathbf{H}(x,y,z,t)}{\partial t},$$

$$\mathbf{H}(x,y,z,t) = s\mathbf{E}(x,y,z,t) + e\frac{\partial \mathbf{E}(x,y,z,t)}{\partial t} + \mathbf{J}(x,y,z,t),$$

$$\nabla \cdot \mathbf{E}(x,y,z,t) = \frac{r_v(x,y,z,t)}{e},$$

$$\nabla \cdot \mathbf{H}(x,y,z,t) = 0,$$

where $\boldsymbol{J}$ is the current density, and using the conductivity s, we have $\mathbf{J} = s\mathbf{E}$; r_v is the volume charge density.

The total current density is the sum of the source current $\mathbf{J}_S$ and the conduction current density $s\mathbf{E}$ (due to the field created by the source $\mathbf{J}_S$). Thus,

$$\mathbf{J} = \mathbf{J}_S + s\mathbf{E}.$$

The equation of conservation of charge (continuity equation) is

$$\oint_s \mathbf{J} \cdot d\mathbf{s} = -\frac{d}{dt}\int_v r_v dv,$$

and in the point form one obtains

$$\nabla \cdot \mathbf{J}(x,y,z,t) = -\frac{r_v(x,y,z,t)}{t}.$$

Therefore, the net outflow of current from a closed surface results in decrease of the charge enclosed by the surface.

The electromagnetic waves transfer the electromagnetic power. That is, the energy is delivered by means of electromagnetic waves. Using equations

$$\mathbf{E} = \mu \frac{\partial \mathbf{H}}{\partial t} \quad \text{and} \quad \mathbf{H} = \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J},$$

we have

$$\nabla \times (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \times (\nabla \times \mathbf{E}) - \mathbf{E} \times (\nabla \times \mathbf{H}) = -\mathbf{H} \times \mu \frac{\partial \mathbf{H}}{\partial t} - \mathbf{E} \times \left(\varepsilon \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \right).$$

In a media, where the constitute parameters are constant (time-invariant), we have the so-called point-function relationship

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\frac{\partial}{\partial t}\left(\tfrac{1}{2}\varepsilon E^2 + \tfrac{1}{2}\mu H^2 \right) - \sigma E^2.$$

In integral form one obtains

$$\oint_s (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{s} = -\frac{\partial}{\partial t} \int_v \left(\tfrac{1}{2}\varepsilon E^2 + \tfrac{1}{2}\mu H^2 \right) dv - \int_v \sigma E^2 dv.$$

time rate of change of energy stored in the electric field $\mathbf{E}$ and magnetic field $\mathbf{H}$

ohmic power dissipated in the presence of $\mathbf{E}$

The right side of the equation derived gives the rate of decrease of the electric and magnetic energies stored minus the ohmic power dissipated as heat in the volume *v*. The pointing vector, which is a power density vector, represents the power flows per unit area, and

$$\mathbf{P} = \mathbf{E} \times \mathbf{H}.$$

Furthermore,

$$\oint_s (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{s} = \oint_s \mathbf{P} \cdot d\mathbf{s} = -\frac{\partial}{\partial t} \int_v (w_E + w_H) dv + \int_v r_s \, dv,$$

power leaving the enclosed volume

where $w_E = \frac{1}{2}\varepsilon E^2$ and $w_H = \frac{1}{2}\mu H^2$ are the electric and magnetic energy densities; $r_s = \sigma E^2 = \frac{1}{\sigma} J^2$ is the ohmic power density.

The important conclusion is that the total power transferred into a closed surface ***s*** at any instant equals the sum of the rate of increase of the stored electric and magnetic energies and the ohmic power dissipated within the enclosed volume *v*.

If the source charge density $r_v(x, y, z, t)$ and the source current density $\mathbf{J}(x, y, z, t)$ vary sinusoidally, the electromagnetic field also vary sinusoidally. Hence, we have deal with the so-called time-harmonic electromagnetic fields. The sinusoidal time-varying electromagnetic fields will be studied. Hence, the phasor analysis is applied. For example,

$$\mathbf{E}(\mathbf{r}) = E_x(\mathbf{r})\mathbf{a}_x + E_y(\mathbf{r})\mathbf{a}_y + E_z(\mathbf{r})\mathbf{a}_z.$$

The electric field intensity components are the complex functions. In particular,

$E_x(\mathbf{r}) = E_{x\mathrm{Re}} + jE_{x\mathrm{Im}}$, $E_y(\mathbf{r}) = E_{y\mathrm{Re}} + jE_{y\mathrm{Im}}$, $E_z(\mathbf{r}) = E_{z\mathrm{Re}} + jE_{z\mathrm{Im}}$.

For the real electromagnetic field, we have

$E_x(\mathbf{r},t) = E_{x\mathrm{Re}}(\mathbf{r})\cos wt \quad E_{x\mathrm{Im}}(\mathbf{r})\sin wt$.

One obtains the time-harmonic electromagnetic field equations. In particular,

Faraday's law $\mathbf{E} = \quad jwm\mathbf{H}$,

generalized Amphere's law $\mathbf{H} = s\mathbf{E} + jwe\mathbf{E} + \mathbf{J} = jw\left(\frac{s}{jw} + e\right)\mathbf{E} + \mathbf{J}$,

Gauss's law $\nabla\cdot\mathbf{E} = \frac{r_v}{\frac{s}{jw} + e}$,

Continuity of magnetic flux $\nabla\cdot\mathbf{H} = 0$,

Continuity law $\nabla\cdot\mathbf{J} = \quad jwr_v$, (4.1.1)

where $\frac{s}{jw} + e$ is the complex permittivity. However, for simplicity we will use e keeping in mind that the expression for the complex permittivity $\frac{s}{jw} + e$ must be applied.

The electric field intensity ***E***, electric flux density ***D***, magnetic field intensity ***H***, magnetic flux density ***B***, and current density **J** are complex-valued functions of spatial coordinates.

From the equation (4.1.1) taking the curl of $\mathbf{E} = \quad jwm\mathbf{H}$, which is rewritten as $\mathbf{E} = \quad jw\mathbf{B}$, and using $\mathbf{H} = jw\mathbf{D} + \mathbf{J}$, one obtains

$$\nabla\times\nabla\times\mathbf{E} = w^2me\mathbf{E} = k_v^2\mathbf{E} = \quad jwm\mathbf{J},$$

where k_v is the wave constant $k_v = w\sqrt{me}$, and in free space $k_{v0} = w\sqrt{m_0e_0} = \frac{w}{c}$ because the speed of light is $c = \frac{1}{\sqrt{m_0e_0}}$, $c = 3 \quad 10^8 \ \frac{\mathrm{m}}{\mathrm{sec}}$.

The wavelength is found as

$$l_v = \frac{2p}{k_v} = \frac{2p}{w\sqrt{me}},$$

and in free space $l_{v0} = \frac{2p}{k_{v0}} = \frac{2pc}{w}$.

Using the magnetic vector potential $\boldsymbol{A}$, we have $\mathbf{B} = \quad \mathbf{A}$.

Hence,

$$(\mathbf{E} + jw\mathbf{A}) = 0,$$

and thus

$$\mathbf{E} + jw\mathbf{A} = \quad \ddot{\mathbf{E}},$$

where $\ddot{\mathbf{E}}$ is the scalar potential.

To guarantee that $\mathbf{H} = jw\mathbf{D} + \mathbf{J}$ holds, it is required that

$$m\mathbf{H} = \quad \mathbf{A} = \quad \times\mathbf{A} \quad {}^2\mathbf{A} = jwme\mathbf{E} + m\mathbf{J}.$$

Therefore, one finally finds the equation needed to be solved

$${}^2\mathbf{A} + k_v^2\mathbf{A} = \quad (\quad \times\mathbf{A} + jwme\ddot{\mathbf{E}}) \quad m\mathbf{J}.$$

Taking note of the Lorentz condition $\times\mathbf{A} = \quad jwme\ddot{\mathbf{E}}$, one obtains.

$${}^2\mathbf{A} + k_v^2\mathbf{A} = \quad m\mathbf{J}.$$

Thus, the equation for $\ddot{\mathbf{E}}$ is found. In particular,

$${}^2\ddot{\mathbf{E}} + k_v^2\ddot{\mathbf{E}} = \quad \frac{r_v}{e}.$$

The equation for the magnetic vector potential is found solving the following inhomogeneous Helmholtz equation

$${}^2\mathbf{A} + k_v^2\mathbf{A} = \quad m\mathbf{J}.$$

The expression for the electromagnetic field intensity, in terms of the vector potential, is

$$\mathbf{E} = \quad jw\mathbf{A} + \frac{\times\mathbf{A}}{jwme}.$$

To derive **E**, one must have **A**. The Laplacian for **A** in different coordinate systems can be found. For example, we have

$${}^2A_x + k_v^2A_x = \quad mJ_x,$$

$${}^2A_y + k_v^2A_y = \quad mJ_y,$$

$${}^2A_z + k_v^2A_z = \quad mJ_z.$$

It was shown that the magnetic vector potential and the scalar potential obey the time-dependent inhomogeneous wave equation

$${}^2 \quad k\frac{\partial^2}{\partial t^2} \quad (\mathbf{r},t) = \quad F(\mathbf{r},t).$$

The solution of this equation is found using Green's function as

$$(\mathbf{r},t) = \qquad F(\mathbf{r}',t')G(\mathbf{r} \quad \mathbf{r}';t \quad t')dt'dt',$$

where $G(\mathbf{r} \quad \mathbf{r}';t \quad t') = \quad \frac{1}{4p|\mathbf{r} \quad \mathbf{r}'|}d\left(t \quad t' \quad k|\mathbf{r} \quad \mathbf{r}'|\right).$

The so-called retarded solution is

$$(\mathbf{r},t) = \int \frac{F(\mathbf{r}', t' - k|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} dt' .$$

For sinusoidal electromagnetic fields, we apply the Fourier analysis to obtain

$$(\mathbf{r}) = \frac{1}{4p} \int \frac{e^{-jk_v|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} F(\mathbf{r}') dt' .$$

Thus, we have the expressions for the phasor retarded potentials

$$\mathbf{A}(\mathbf{r}) = \frac{m}{4p} \int_v \frac{e^{-jk_v|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \mathbf{J}(\mathbf{r}') dv ,$$

$$\ddot{\mathbf{E}}(\mathbf{r}) = \frac{1}{4pe} \int_v \frac{e^{-jk_v|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} r(\mathbf{r}') dv .$$

Example 4.1.1.

Consider a short (*dl*) thin filament of current located in the origin, see Figure 4.1.1. Derive the expressions for magnetic vector potential and electromagnetic field intensities.

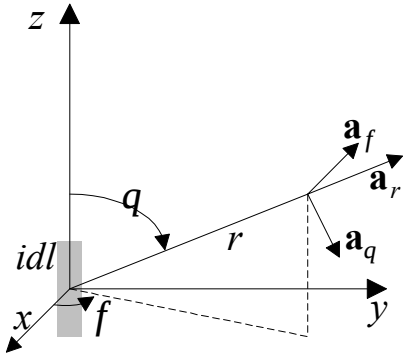


Figure 4.1.1. Current filament in the spherical coordinate system

Solution.

The magnetic vector potential has only a *z* component, and thus, from

$$\nabla^2 \mathbf{A} + k_v^2 \mathbf{A} = -m\mathbf{J} ,$$

we have

$$\nabla^2 A_z + k_v^2 A_z = -mJ_z = -m\frac{i}{ds} ,$$

where *ds* is the cross-sectional area of the filament.

Taking note of the spherical symmetry, we conclude that the magnetic vector potential A_z is not a function of the polar and azimuth angles q and f. In particular, the following equation results

$$\frac{1}{r^2}\frac{}{r}r^2\frac{A_z}{r}+k_v^2A_z=0.$$

It is well-known that the solution of equation $\frac{d^2y}{dy^2}+k_v^2y=0$ has two components. In particular, e^{jk_vr} (outward propagation) and e^{jk_vr} (inward propagation). The inward propagation is not a part of solution for the filament located in the origin. Thus, we have

$y(t,r)=ae^{jwt\ jk_vr}$ (outward propagating spherical wave).

In free space, we have

$y(t,r)=ae^{jw(t\ r/c)}$.

Substituting $A_z=\frac{y}{r}$, one obtains $A_z(r)=\frac{a}{r}e^{\frac{jwr}{c}}$.

To find the constant *a*, we use the volume integral

$$^2A_zdv=\circ\ \ A_z\times\mathbf{a}_rr_d^2\sin qdqdf=\ \frac{w^2}{c^2}\ A_zdv\ \ m_0\ J_zdv,$$

where the differential spherical volume is $dv=r_d^2\sin qdqdfdr$; r_d is the differential radius.

Making use of

$$A_z\times\mathbf{a}_r=\frac{A_z}{r}=\ \ 1+j\frac{w}{c}r\div\frac{a}{r^2}e^{\ j\frac{w}{c}r},$$

we have

$$\lim_{r_d\ 0}^{2pp}\ \ 1+j\frac{w}{c}r_d\div ae^{\ j\frac{w}{c}r_d}\sin qdqdf=\ 4pa=\ m_0idl,$$

one has $a=\frac{m_0idl}{4p}$.

Thus, the following expression results

$$A_z(r)=\frac{m_0idl}{4pr}e^{\ \frac{jwr}{c}}.$$

Therefore, the final equation for the magnetic vector potential (outward propagating spherical wave) is

$$\mathbf{A}(r)=\frac{m_0idl}{4pr}e^{\ \frac{jwr}{c}}\mathbf{a}_z.$$

From $\mathbf{a}_z = \mathbf{a}_r \cos q \quad \mathbf{a}_q \sin q$, we have

$$\mathbf{A}(r) = \frac{m_0 idl}{4pr} e^{\frac{jwr}{c}} (\mathbf{a}_r \cos q \quad \mathbf{a}_q \sin q)$$

The magnetic and electric field intensities are found using

$$\mathbf{B} = \quad \mathbf{A} \text{ and } \mathbf{E} = \quad jw\mathbf{A} + \frac{\times \mathbf{A}}{jwme}.$$

Then, one finds

$$\mathbf{H}(r) = \frac{1}{m_0} \qquad \mathbf{A}(r) = \frac{idl \sin q}{4pr} \quad \frac{jw}{cr} + \frac{1}{r^2} \div e^{\frac{jwr}{c}} \mathbf{a}_f ,$$

$$\mathbf{E}(r) = \frac{j\sqrt{\frac{m_0}{e_0}} cidl}{4pw} \cos q \quad \frac{jw}{cr^2} + \frac{1}{r^3} \div e^{\frac{jwr}{c}} \mathbf{a}_r$$

$$\frac{j\sqrt{\frac{m_0}{e_0}} cidl}{4pw} \sin q \qquad \frac{w^2}{c^2 r} + \frac{jw}{cr^2} + \frac{1}{r^3} \div e^{\frac{jwr}{c}} \mathbf{a}_q .$$

The intrinsic impedance is given as

$$Z_0 = \sqrt{\frac{m_0}{e_0}} \text{ , and } Y_0 = \frac{1}{Z_0} = \sqrt{\frac{e_0}{m_0}} .$$

Near-field and far-field electromagnetic radiation fields can be found, simplifying the expressions for **H**(*r*) and **E**(*r*).

For near-field, we have

$$\mathbf{H}(r) = \frac{1}{m_0} \qquad \mathbf{A}(r) = j \frac{idl \sin qw}{4pcr^2} e^{\frac{jwr}{c}} \mathbf{a}_f$$

and $\mathbf{E}(r) = j \dfrac{\sqrt{\frac{m_0}{e_0}} cidlw}{4pc^2 r} \sin q \mathbf{a}_q .$

The complex Pointing vector can be found as

$\frac{1}{2} \mathbf{E}(r) \quad \mathbf{H}^*(r)$.

The following expression for the complex power flowing out of a sphere of radius *r* results

$$\frac{1}{2} \circ_s \left(\mathbf{E}(r) \quad \mathbf{H}^*(r) \right) \times d\mathbf{s} = \frac{w^2 m_0 \sqrt{m_0 e_0} i^2 dl^2}{12p} = \frac{w m_0 k_v i^2 dl^2}{12p} .$$

The real quality is found, and the power dissipated in the sense that it travels away from source and cannot be recovered.

Example 4.1.2.

Derive the expressions for the magnetic vector potential and electromagnetic field intensities for a magnetic dipole (small current loop) which is shown in Figure 4.1.2.

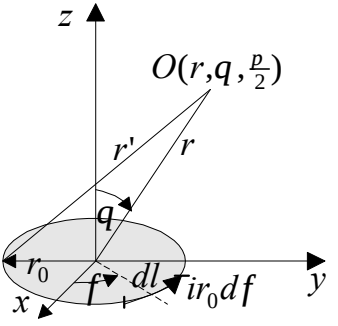


Figure 4.1.2. Current loop in the *xy* plane

Solution.

The magnetic dipole moment is equal to the current loop are times current. That is,

$$\mathbf{M} = p r_0^2 i \mathbf{a}_z = M \mathbf{a}_z .$$

For the short current filament, it was derived in Example 4.1.1 that

$$\mathbf{A}(r) = \frac{m_0 i dl}{4pr} e^{\frac{jwr}{c}} \mathbf{a}_z .$$

In contrast, we have

$$\mathbf{A} = \frac{m_0 i}{4p} \circ_l \frac{1}{r'} dl .$$

The distance between the source element *dl* and point $O(r,q,\frac{p}{2})$ is denoted as *r*'. It should be emphasized that the current filament is lies in the *xy* plane, and

$$dl = \mathbf{a}_f r_0 df = (\ \mathbf{a}_x \sin f + \mathbf{a}_y \cos f) r_0 df .$$

Thus, due to the symmetry

$$\mathbf{A} = \mathbf{a}_f \frac{m_0 i r_0}{2p} \int_{-p/2}^{p/2} \frac{\sin f}{r'} df ,$$

where using the trigonometric identities one finds

$$r'^2 = r^2 + r_0^2 \quad 2 r r_0 \sin q \sin f .$$

Assuming that $r^2 >> r_0^2$, we have $\frac{1}{r'} \quad \frac{1}{r} \ 1 + \frac{r_0}{r} \sin q \sin f \div .$

Therefore

$$\mathbf{A} = \mathbf{a}_f \frac{m_0 i r_0}{2p} \int_{-p/2}^{p/2} \frac{\sin f}{r'} df$$

$$= \mathbf{a}_f \frac{m_0 i r_0}{2pr} \int_{-p/2}^{p/2} \left(1 + \frac{r_0}{r} \sin q \sin f\right) \sin f df = \mathbf{a}_f \frac{m_0 i r_0^2}{4r^2} \sin q .$$

Having obtained the explicit expression for the vector potential, the magnetic flux density is found. In particular,

$$\mathbf{B} = \nabla \times \mathbf{A} = \nabla \times \mathbf{a}_f \frac{m_0 i r_0^2}{4r^2} \sin q = \frac{m_0 i r_0^2}{4r^3} (2\mathbf{a}_r \cos q + \mathbf{a}_q \sin q) .$$

Taking note of the expression for the magnetic dipole moment $\mathbf{M} = p r_0^2 i \mathbf{a}_z$, one has

$$\mathbf{A} = \mathbf{a}_f \frac{m_0 i r_0^2}{4r^2} \sin q = \frac{m_0}{4p r^2} \mathbf{M} \times \mathbf{a}_r .$$

It was shown that using $\mathbf{A} = \frac{m_0 i}{4p} \oint_l \frac{1}{r'} dl$, the desired results are obtained.

Let us apply $\mathbf{A} = \frac{m_0 i}{4p} \oint_l \frac{e^{-j\frac{w}{c} r'}}{r'} dl$.

From $e^{-j\frac{w}{c} r'} \approx 1 - j\frac{w}{c}(r' - r)$ $e^{-j\frac{w}{c} r}$, we have

$$\mathbf{A} = \frac{m_0 i}{4p} \oint_l \frac{[1 - j\frac{w}{c}(r' - r)] e^{-j\frac{w}{c} r}}{r'} dl = \mathbf{a}_f \frac{m_0 M}{4p r^2} \left(1 + j\frac{w}{c} r\right) e^{-j\frac{w}{c} r} \sin q .$$

Therefore, one finds

$$E_f = j \frac{m_0 w^3 M}{4c^2 p} \frac{1}{j\frac{w}{c} r} \frac{1}{\frac{w^2}{c^2} r^2} \div e^{-j\frac{w}{c} r} \sin q ,$$

$$H_r = j \frac{2 m_0 w^3 M}{4c^2 \sqrt{\frac{m_0}{e_0}} p} \left(\frac{1}{\frac{w^2}{c^2} r^2} + \frac{1}{j\frac{w^3}{c^3} r^3}\right) \div e^{-j\frac{w}{c} r} \cos q ,$$

$$H_q = -j \frac{m_0 w^3 M}{4c^2 \sqrt{\frac{m_0}{e_0}} p} \frac{1}{j\frac{w}{c} r} \left(\frac{1}{\frac{w^2}{c^2} r^2} - \frac{1}{j\frac{w^3}{c^3} r^3}\right) \div e^{-j\frac{w}{c} r} \sin q .$$

The electromagnetic fields in near- and far-fields can be straightforwardly derived, and thus, the corresponding approximations for the E_f, H_r and H_q can be obtained.

Let the current density distribution in the volume is given as $\mathbf{J}(\mathbf{r}_0)$, and for far-field from Figure 4.1.3 one has $\mathbf{r}$ $\mathbf{r}'$ $\mathbf{r}_0$.

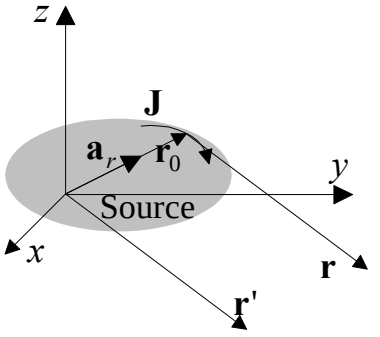


Figure 4.1.3. Radiation from volume current distribution

The formula to calculate far-field magnetic vector potential is

$$\mathbf{A}(\mathbf{r}) = \frac{m}{4pr} e^{\ jk_v r} \int_v \mathbf{J}(\mathbf{r}_0) e^{\ jk_v \mathbf{r}_0} dv ,$$

and the electric and magnetic field intensities are found using

$$\mathbf{E} = \ jw\mathbf{A} + \frac{\times\mathbf{A}}{jwme}$$

and $\mathbf{B} =$ $\mathbf{A}$.

We have

$$\mathbf{E}(\mathbf{r}) = \frac{jk_v Z_v}{4pr} e^{\ jk_v r} \int_v \left[\mathbf{a}_r \times \mathbf{J}(\mathbf{r}_0)\mathbf{a}_r \quad \mathbf{J}(\mathbf{r}_0)\right] e^{\ jk_v \mathbf{r}_0} dv ,$$

$$\mathbf{H}(\mathbf{r}) = Y_v \mathbf{a}_r \quad \mathbf{E}(\mathbf{r}) .$$

Example 4.1.3.

Consider the half-wave dipole antenna fed from a two-wire transmission line, as shown in Figure 4.1.4 The antenna is one-quarter wavelength; that is, $\frac{1}{4} l_v$ z $\frac{1}{4} l_v$. The current distribution is $i(z) = i_0 \cos k_v z$. Obtain the equations for electromagnetic field intensities and radiated power.

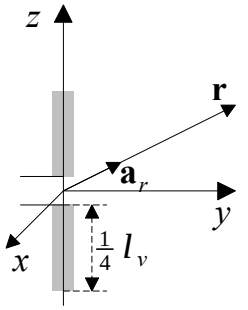


Figure 4.1.4. Half-wave dipole antenna

Solution.

The wavelength is given as $l_v = \dfrac{2p}{k_v} = \dfrac{2p}{w\sqrt{me}}$, and in free space we have $l_{v0} = \dfrac{2p}{k_{v0}} = \dfrac{2pc}{w}$.

It was emphasized that and $k_{v0} = w\sqrt{m_0 e_0}$.

Making use of

$$\mathbf{E}(\mathbf{r}) = \frac{jk_v Z_v}{4pr} e^{-jk_v r} \int_v \left[\mathbf{a}_r \times \mathbf{J}(\mathbf{r}_0)\mathbf{a}_r \quad \mathbf{J}(\mathbf{r}_0)\right] e^{jk_v \mathbf{r}_0} dv ,$$

we have the following line integral

$$\mathbf{E}(\mathbf{r}) = \frac{jk_v Z_v}{4pr} e^{-jk_v r} \circ \int_l \left[(\mathbf{a}_r \times \mathbf{a}_l)\mathbf{a}_r \quad \mathbf{a}_l\right] i(l) e^{jk_v \mathbf{r}_0} dl ,$$

where $\mathbf{a}_l$ is the unit vector in the current direction.

Then,

$$\mathbf{E}(\mathbf{r}) = \frac{jk_v Z_v i_0}{4pr} e^{-jk_v r} \int_{-\frac{1}{4}l_v}^{\frac{1}{4}l_v} (\mathbf{a}_r \cos q \quad \mathbf{a}_z) \cos k_v z\, e^{jk_v z \cos q} dz$$

$$= \frac{jZ_v i_0 \cos(\frac{1}{2}p\cos q)}{2pr \sin q} e^{-jk_v r} \mathbf{a}_q .$$

Having found the magnetic field intensity as

$$\mathbf{H}(\mathbf{r}) = Y_v \mathbf{a}_r \times \mathbf{E}(\mathbf{r}) = H_f \mathbf{a}_f = \frac{ji_0 \cos(\frac{1}{2}p\cos q)}{2pr \sin q} e^{-jk_v r} \mathbf{a}_f ,$$

the power flux per unit area is

$$\tfrac{1}{2}\mathrm{Re}\left(\mathbf{E}(\mathbf{r}) \times \mathbf{H}(\mathbf{r})^* \cdot \mathbf{a}_r\right) = \tfrac{1}{2} E_f H_f^* = \frac{|i_0|^2 Z_0 \cos^2(\frac{1}{2}p\cos q)}{8p^2 r^2 \sin^2 q} ,$$

and integrating the derived expression over the surface $\frac{|i_0|^2 Z_0}{8p^2}\int_0^{2p}\int_0^{p}\frac{\cos^2(\frac{1}{2}p\cos q)}{\sin^2 q}\sin q\,dq\,df$, the total radiated power is found to be $36.6|i_0|^2$.

If the current density distribution is known, the radiation field can be found. Using Maxwell's equations, using the electric and magnetic vector potentials $\mathbf{A}_E$ and $\mathbf{A}_H$, we have the following equations

$$\left(\nabla^2 + k_v^2\right)\mathbf{A}_E = -m\mathbf{J}_E,$$

$$\left(\nabla^2 + k_v^2\right)\mathbf{A}_H = -e\mathbf{J}_H,$$

$$\mathbf{E} = -jw\mathbf{A}_E + \frac{1}{jwme}\nabla\times\mathbf{A}_E - \frac{1}{e}\mathbf{A}_H,$$

$$\mathbf{H} = -jw\mathbf{A}_H + \frac{1}{jwme}\nabla\times\mathbf{A}_H - \frac{1}{m}\mathbf{A}_E.$$

The solutions are

$$\mathbf{A}_E(\mathbf{r}) = \frac{m}{4pr}e^{-jk_v r}\int_v e^{jk_v\mathbf{r}}\mathbf{J}_E(\mathbf{r})d\mathbf{r},$$

$$\mathbf{A}_H(\mathbf{r}) = \frac{e}{4pr}e^{-jk_v r}\int_v e^{jk_v\mathbf{r}}\mathbf{J}_H(\mathbf{r})d\mathbf{r}.$$

Example 4.1.4.

Consider the slot (one-half wavelength long slot is dual to the half-wave dipole antenna studied in Example 4.1.3), which is exited from the coaxial line, see Figure 4.1.5. The electric field intensity in the *z*-direction is $E = E_0\sin k_v(l - |z|)$. Derive the expressions for the magnetic vector potential and electromagnetic field intensities.

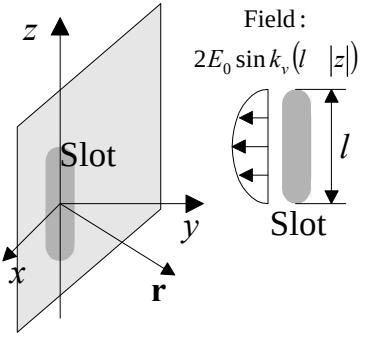


Figure 4.1.5. Slot antenna

Solution.

Using the magnetic current density $\mathbf{J}_H$, from

$$\oint_s \mathbf{E}\times d\mathbf{s} = \oint_l \mathbf{E}\times d\mathbf{l} = -jw\int_s \mathbf{B}\times d\mathbf{s} = -\int_s \mathbf{J}_H \times d\mathbf{s},$$

the boundary conditions for the magnetic current sheet are found as

$$\mathbf{a}_n \quad \mathbf{E}_1 \quad \mathbf{a}_n \quad \mathbf{E}_2 = -\mathbf{J}_H .$$

The slot antenna is exited by the magnetic current with strength $2E_0 \sin k_v\left(l - |z|\right)$ in the z axis. For half-wave slot we have $i_H = i_0 \sin k_v\left(l - |z|\right)$, and

$$\left(\nabla^2 + k_v^2\right)\mathbf{A}_H = -e\mathbf{J}_H ,$$

$$\mathbf{H} = -jw\mathbf{A}_H + \frac{1}{jwme}\nabla\times\mathbf{A}_H = j\frac{i_0 Y_0 \cos(\frac{1}{2}p\cos q)}{2pr\sin q}e^{-jk_v r}\mathbf{a}_q ,$$

$$\mathbf{E} = -\frac{1}{e}\mathbf{A}_H = -j\frac{i_0 \cos(\frac{1}{2}p\cos q)}{2pr\sin q}e^{-jk_v r}\mathbf{a}_f ,,$$

$$\mathbf{A}_H(\mathbf{r}) = \frac{e}{16p}e^{-jk_v r}\int_s e^{jk_v \mathbf{r}}\mathbf{J}(\mathbf{r})d\mathbf{s} .$$

The boundary condition $\mathbf{a}_n \quad \mathbf{E} = -\frac{1}{2}\mathbf{J}_H = \mathbf{a}_n \quad \mathbf{a}_x E_0 \sin k_v\left(l - |z|\right)$ is satisfied by the radiated electromagnetic field.

The radiation pattern of the slot antenna is the same as for the dipole antenna.

References

1. Hayt W. H., *Engineering Electromagnetics*, McGraw-Hill, New York, 1989.
2. Collin R. E., *Antennas and Radiowave propagation*," McGraw-Hill, New York, 1985.
3. Paul C. R., Whites K. W., and Nasar S. A., *Introduction to Electromagnetic Fields*, McGraw-Hill, New York, 1998.

4.2. DESIGN OF CLOSED-LOOP NANO- AND MICROELECTROMECHANICAL SYSTEMS USING THE LYAPUNOV STABILITY THEORY

The solution of a spectrum of problems in nonlinear analysis, structural synthesis, modeling, and optimization of NEMS and MEMS lead to the development of superior high-performance NEMS and MEMS. In this section, we address introductory control issues. Mathematical models of NEMS and MEMS were derived, and the application of the Lyapunov theory is studied as applied to solve the motion control problem.

It was illustrated that NEMS and MEMS must be controlled. Nano- and microelectromechanical systems augment a great number of subsystems, and to control microscale electric motors, as discussed in previous chapters, power amplifiers (ICs) regulate the voltage or current fed to the motor windings. These power amplifiers are controlled based upon the reference (command), output, decision making, and other variables. Studying the end-to-end NEMS and MEMS behavior, usually the output is the nano- or microactuator linear and angular displacements. There exist infinite number of possible NEMS and MEMS configurations, and it is impossible to cover all possible scenarios. Therefore, our efforts will be concentrated on the generic results which can be obtained describing NEMS and MEMS by differential equations. That is, using the mathematical model, as given by differential equations, our goal is develop control algorithms to guarantee the desired performance characteristics addressing the motion control problem (settling time, accuracy, overshoot, controllability, stability, disturbance attenuation, et cetera).

Several methods have been developed to address and solve nonlinear design and motion control problems for multi-input/multi-output dynamic systems. In particular, the Hamilton-Jacobi and Lyapunov theories are found to be the most straightforward in the design of control laws.

The NEMS and MEMS dynamics is described as

$$\dot{x}(t) = F(x,r,d) + B(x)u \, , \, y = H(x) \, , \, u_{\min} \quad u \quad u_{\max} \, , x(t_0) = x_0 \, , \tag{4.2.1}$$

where $x \;\; X \;\; |^{c}$ is the state vector; $u \;\; U \;\; |^{m}$ is the bounded control vector; $r \;\; R \;\; |^{b}$ and y $\;\; Y \;\; |^{b}$ are the measured reference and output vectors; $d \;\; D \;\; |^{s}$ is the disturbance vector; $F(\cdot): |^{c} \;\; |^{b} \;\; |^{s} \;\; |^{c}$ and $B(\cdot): |^{c} \;\; |^{c \times m}$ are jointly continuous and Lipschitz; $H(\cdot): |^{c} \;\; |^{b}$ is the smooth map defined in the neighborhood of the origin, $H(0) = 0$.

Before engaged in the design of closed-loop systems, which will be based upon the Lyapunov stability theory, let us study stability of time-varying nonlinear dynamic systems described by

$$\dot{x}(t) = F(t,x) \, , \;\; x(t_0) = x_0 \, , \;\; t \quad 0 \, .$$

The following Theorem is formulated.

Theorem.

Consider the system described by nonlinear differential equations

$\dot{x}(t) = F(t,x)$, $x(t_0) = x_0$, $t \quad 0$.

If there exists a positive-definite scalar function $V(t,x)$ *(called Lyapunov function) with continuous first-order partial derivatives with respect to t and x*

$$\frac{dV}{dt} = \frac{V}{t} + \quad \frac{V}{x} \div \frac{dx}{dt} = \frac{V}{t} + \quad \frac{V}{x} \div F(t,x),$$

then

the equilibrium state is stable if the total derivative of the positive-definite function $V(t,x) > 0$ *is* $\frac{dV}{dt} \quad 0$ *;*

the equilibrium state is uniformly stable if the total derivative of the positive-definite decreasing function $V(t,x) > 0$ *is* $\frac{dV}{dt} \quad 0$ *;*

the equilibrium state is uniformly asymptotically stable in the large if the total derivative of $V(t,x) > 0$ *is negative definite; that is,* $\frac{dV}{dt} < 0$ *;*

the equilibrium state is exponentially stable in the large if the exist the K *-functions* $r_1(\cdot)$ *and* $r_2(\cdot)$*, and K-function* $r_3(\cdot)$ *such that*

$r_1(\| x \|) \quad V(t,x) \quad r_2(\| x \|)$ *and* $\frac{dV(x)}{dt} \qquad r_3(\| x \|)$. ■

Examples are studied to illustrate how this Theorem can be straightforwardly applied.

Example 4.2.1.

Study stability of the following nonlinear system

$\dot{x}_1(t) = \quad x_1^3 \quad x_1 x_2^2$,

$\dot{x}_2(t) = \quad x_2^7$, $t \quad 0$.

Solution.

A scalar positive-definite function is expressed in the quadratic form as

$V(x_1, x_2) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$.

The total derivative is found to be

$$\frac{dV(x_1,x_2)}{dt} = \frac{V}{x} \div \frac{dx}{dt} = \frac{V}{x} \div F(x)$$

$$= \frac{V}{x_1}\left(x_1^3 \quad x_1 x_2^2 \right) + \frac{V}{x_2}\left(x_2^7 \right) = x_1^4 \quad x_1^2 x_2^2 \quad x_2^8 .$$

Thus, we have $\frac{dV(x_1,x_2)}{dt} < 0$.

The total derivative of $V(x_1,x_2) > 0$ is negative definite.

Therefore, the equilibrium state is uniformly asymptotically stable.

Example 4.2.2.

Study stability of the time-varying nonlinear system modeled by the following differential equations

$$\dot{x}_1(t) = x_1 + x_2^3,$$

$$\dot{x}_2(t) = e^{10t} x_1 x_2^2 \quad 5x_2 \quad x_2^3, \; t \quad 0 .$$

Solution.

A scalar positive-definite function is

$V(t,x_1,x_2) = \frac{1}{2}\left(x_1^2 + e^{10t} x_2^2\right)$, $V(t,x_1,x_2) > 0$.

Then, the total derivative, which is expressed as

$$\frac{dV(t,x_1,x_2)}{dt} == \frac{V}{t} + \frac{V}{x_1}\left(x_1 + x_2^3 \right) + \frac{V}{x_2}\left(e^{10t} x_1 x_2^2 \quad 5x_2 \quad x_2^3 \right)$$

$$= x_1^2 \quad e^{10t} x_2^4,$$

is negative definite. In particular, $\frac{dV(x_1,x_2)}{dt} < 0$.

Hence, the equilibrium state is uniformly asymptotically stable.

Example 4.2.3.

Study stability of the systems

$$\dot{x}_1(t) = x_1 + x_2 ,$$

$$\dot{x}_2(t) = x_1 \quad x_2 \quad x_2 |x_2| , \; t \quad 0 .$$

Solution.

The positive-definite scalar Lyapunov candidate is chosen as

$V(x_1,x_2) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$.

Thus, $V(x_1,x_2) > 0$.

The total derivative is

$$\frac{dV(x_1,x_2)}{dt} = x_1\dot{x}_1 + x_2\dot{x}_2 = \quad x_1^2 \quad x_2^2\left(1+\left|x_2\right|\right).$$

Therefore, $\frac{dV(x_1,x_2)}{dt} < 0$. Hence, the equilibrium state is uniformly asymptotically stable, and the quadratic function $V(x_1,x_2) = \frac{1}{2}\left(x_1^2 + x_2^2\right)$ is the Lyapunov function which can be used to study stability.

Example 4.2.4.

Consider a microdrive actuated by permanent-magnet synchronous motor if T_L=0. In drive applications, using equations (3.5.12), three nonlinear differential equations in the rotor reference frame are

$$\frac{di_{qs}^r}{dt} = \quad \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} i_{qs}^r \quad \frac{y_m}{L_{ls} + \frac{3}{2}\overline{L}_m} w_r \quad i_{ds}^r w_r + \frac{1}{L_{ls} + \frac{3}{2}\overline{L}_m} u_{qs}^r,$$

$$\frac{di_{ds}^r}{dt} = \quad \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} i_{ds}^r + i_{qs}^r w_r + \frac{1}{L_{ls} + \frac{3}{2}\overline{L}_m} u_{ds}^r,$$

$$\frac{dw_r}{dt} = \frac{3P^2 y_m}{8J} i_{qs}^r \quad \frac{B_m}{J} w_r.$$

Study the stability letting

1. $u_{qs}^r = 0$ and $u_{ds}^r = 0$ (open-loop system),
2. $u_{qs}^r \quad 0$, $u_{qs}^r = \quad k_w w_r$ and $u_{ds}^r = 0$ (closed-loop system).

Solution.

For open-loop system we have $u_{qs}^r = 0$ and $u_{ds}^r = 0$. Hence,

$$\frac{di_{qs}^r}{dt} = \quad \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} i_{qs}^r \quad \frac{y_m}{L_{ls} + \frac{3}{2}\overline{L}_m} w_r \quad i_{ds}^r w_r,$$

$$\frac{di_{ds}^r}{dt} = \quad \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} i_{ds}^r + i_{qs}^r w_r,$$

$$\frac{dw_r}{dt} = \frac{3P^2 y_m}{8J} i_{qs}^r \quad \frac{B_m}{J} w_r.$$

In matrix form, one obtains

$$\dot{x}(t) = \begin{matrix} \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} & 0 & \frac{y_m}{L_{ls} + \frac{3}{2}\overline{L}_m} \\ 0 & \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} & 0 \\ \frac{3P^2 y_m}{8J} & 0 & \frac{B_m}{J} \end{matrix} \begin{matrix} i_{qs}^r \\ i_{ds}^r \\ w_r \end{matrix} + \begin{matrix} i_{ds}^r w_r \\ i_{qs}^r w_r \\ 0 \end{matrix}.$$

Using the quadratic positive-definite Lyapunov function

$$V(i_{qs}^r, i_{ds}^r, w_r) = \tfrac{1}{2}(i_{qs}^{r\,2} + i_{ds}^{r\,2} + w_r^2),$$

the expression for the total derivative is found to be

$$\frac{dV(i_{qs}^r, i_{ds}^r, w_r)}{dt} =$$

$$\frac{r_s}{L_{ss}}\left(i_{qs}^{r\,2} + i_{ds}^{r\,2}\right) \div \frac{B_m}{J} w_r^{\,2} \quad \frac{8Jy_m - 3P^2 L_{ss} y_m}{8JL_{ss}} i_{qs}^r w_r.$$

Thus,

$$\frac{dV\left(i_{qs}^r, i_{ds}^r, w_r\right)}{dt} < 0.$$

One concludes that the equilibrium state of a microdrive is uniformly asymptotically stable.

Consider the closed-loop system.

To guarantee the balanced operation we let

$$u_{qs}^r = k_w w_r \text{ and } u_{ds}^r = 0.$$

Therefore, the following differential equations result

$$\frac{di_{qs}^r}{dt} = \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} i_{qs}^r \quad \frac{y_m}{L_{ls} + \frac{3}{2}\overline{L}_m} w_r \quad i_{ds}^r w_r \quad \frac{1}{L_{ls} + \frac{3}{2}\overline{L}_m} k_w w_r,$$

$$\frac{di_{ds}^r}{dt} = \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} i_{ds}^r + i_{qs}^r w_r,$$

$$\frac{dw_r}{dt} = \frac{3P^2 y_m}{8J} i_{qs}^r \quad \frac{B_m}{J} w_r,$$

or

$$\dot{x}(t) = \begin{matrix} \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} & 0 & \frac{y_m + k_w}{L_{ls} + \frac{3}{2}\overline{L}_m} \\ 0 & \frac{r_s}{L_{ls} + \frac{3}{2}\overline{L}_m} & 0 \\ \frac{3P^2 y_m}{8J} & 0 & \frac{B_m}{J} \end{matrix} \begin{matrix} i_{qs}^r \\ i_{ds}^r \\ w_r \end{matrix} + \begin{matrix} i_{ds}^r w_r \\ i_{qs}^r w_r \\ 0 \end{matrix}.$$

Taking note of the quadratic positive-definite Lyapunov function

$$V(i_{qs}^r, i_{ds}^r, w_r) = \frac{1}{2}(i_{qs}^{r\ 2} + i_{ds}^{r\ 2} + w_r^2),$$

one has

$$\frac{dV(i_{qs}^r, i_{ds}^r, w_r)}{dt} =$$

$$\frac{r_s}{L_{ss}} \quad i_{qs}^{r\ 2} + i_{ds}^{r\ 2} \div \quad \frac{B_m}{J} w_r^{\ 2} \quad \frac{8J(y_m + k_w) \quad 3P^2 L_{ss} y_m}{8JL_{ss}} i_{qs}^r w_r.$$

Hence, $V\left(i_{qs}^r, i_{ds}^r, w_r\right) > 0$ and $\dfrac{dV\left(i_{qs}^r, i_{ds}^r, w_r\right)}{dt} < 0$.

Therefore, the conditions for asymptotic stability are guaranteed.

In Example 4.2.4 it was shown that dynamic systems can be controlled to attain the desired transient dynamics, stability margins, etc. Let us study how to solve the motion control problem with the ultimate goal to synthesize tracking controllers applying Lyapunov's stability theory.

Using the reference (command) vector $r(t)$ and the system output $y(t)$, the tracking error (which ideally must be zero) is

$$e(t) = Nr(t) \quad y(t). \tag{4.2.2}$$

The Lyapunov theory is applied to derive the *admissible* control laws (voltages and currents are bounded, and therefore the saturation effect is always the reality). That is, the *admissible* bounded controller should be designed as continuous function within the constrained rectangular control set

$U=\{u \quad |^m : u_{\min} \quad u \quad u_{\max}, u_{\min} < 0, u_{\max} > 0\} \quad |^m$.

Making use of the Lyapunov candidate $V(t,x,e)$, the bounded proportional-integral controller with the state feedback extension is expressed as

$$u = \text{sat}_{u_{\min}}^{u_{\max}} \quad G_x(t)B(x)^T \frac{\P V(t,x,e)}{\P x} + G_e(t)B_e^T \frac{\P V(t,x,e)}{\P e} + G_i(t)B_e^T \frac{1}{s}\frac{\P V(t,x,e)}{\P e} \div \tag{4.2.3}$$

where $G_x(\cdot): |_0 \quad |^{m\ m}$, $G_e(\cdot): |_0 \quad |^{m\ m}$ and $G_i(\cdot): |_0 \quad |^{m\ m}$ are the bounded symmetric matrix-functions defined on $[t_0, \ )$, $G_x>0$, $G_e>0$, $G_i>0$; $V(\cdot): |_0 \quad |^c \quad |^b \quad |_0$ is the continuously differentiable real-analytic C^k $(k \ 1)$ function with respect to $x \ X$ and $e \ E$ on $[t_0, \ )$.

It was emphasized that the control signal is saturated as documented in Figure 4.2.1.

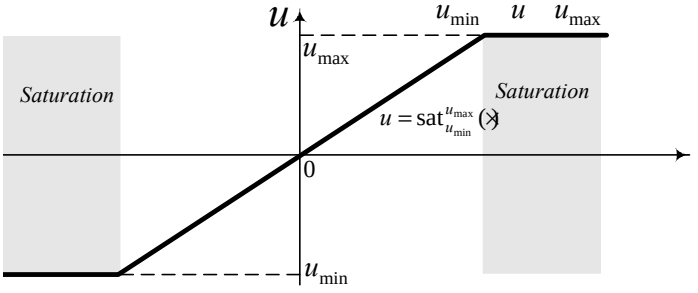


Figure 4.2.1. Bounded control, $u_{\min} \quad u \quad u_{\max}$

For closed-loop NEMS and MEMS (4.2.1)–(4.2.3) with $X_0=\{x_0 \ |_c\} \phi X \ |_c$, $u \ U \ |_m$, $r \ R \ |_b$ and $d \ D \ |_s$, it is straightforward to find the evolution set $X(X_0, U, R, D) \ |^c$. Furthermore, using the output equation, one has $X \quad ^H \quad Y$. Thus, the system (4.2.1)-(4.2.3) evolves in $XY(X_0,U, R, D)=\{(x,y) \ X \ Y: x_0 \ X_0, u \ U, r \ R, d \ D, t \ [t_0, \ )\} \ |^c \ |^b$.

The tracking error

$$e(t) = Nr(t) \quad y(t), e(\cdot):[t_0, \ ) \ |^b$$

gives the difference between the reference input $r(\cdot):[t_0, \ ) \ |^b$ and system output $y(\cdot):[t_0, \ ) \ |^b$. Our goal is to find the feedback coefficients of controller (4.2.3) to guarantee that the closed-loop NEMS and MEMS will evolve in the desired manner. The following Lyapunov-based Lemma is formulated to study the stability of closed-loop dynamic systems as well as to find the feedback coefficients to guarantee the criteria imposed on the Lyapunov pair.

Lemma.

Consider the closed-loop systems (4.2.1) – (4.2.3).

1. *Solutions of system are uniformly ultimately bounded;*
2. *equilibrium point is exponentially stable in the convex and compact state evolution set $X(X_0, U, R, D) \ |^c$;*
3. *tracking is ensured and disturbance attenuation is guaranteed in the state-error evolution set $XE(X_0, E_0, U, R, D) \ |^c \ |^b$,*

if there exists a C^k function $V(t,x,e)$ in XE such that for all $x \ X, e \ E, u \ \hat{I} \ U, r \ R$ and $d \ D$ on $[t_0, \ )$

(i) $$r_1 \| x \| + r_2 \| e \| \quad V(t,x,e) \quad r_3 \| x \| + r_4 \| e \|, \tag{4.2.4}$$

(ii) along (4.2.1) with (4.2.3), the following inequality holds

$$\frac{dV(t,x,e)}{dt} \quad r_5 \| x \| \quad r_6 \| e \|. \tag{4.2.5}$$

Here, $_1(\cdot): |_0 \ |_0$, $_2(\cdot): |_0 \ |_0$, $_3(\cdot): |_0 \ |_0$ and $_4(\cdot): |_0 \ |_0$ are the K -functions; $_5(\cdot): |_0 \ |_0$ and $_6(\cdot): |_0 \ |_0$ are the K-functions.

The major problem is to design the Lyapunov candidate functions.

Let us apply a family of nonquadratic Lyapunov candidates

$$V(t,x,e)=\sum_{i=0}^{h}\frac{2g+1}{2(i+g+1)}\,x^{\frac{i+g+1}{2g+1}\,T}\div K_{xi}(t)x^{\frac{i+g+1}{2g+1}}+\sum_{i=0}^{V}\frac{2b+1}{2(i+b+1)}\,e^{\frac{i+b+1}{2b+1}\,T}\div K_{ei}(t)e^{\frac{i+b+1}{2b+1}}$$
$$+\sum_{i=0}^{s}\frac{2m+1}{2(i+m+1)}\,e^{\frac{i+m+1}{2m+1}\,T}\div K_{si}(t)e^{\frac{i+m+1}{2m+1}}. \qquad (4.2.6)$$

To design the Lyapunov functions, the nonnegative integers were used.

In particular, $h=0,1,2,...$, $g=0,1,2,...$, $V=0,1,2,...$, $b=0,1,2...$, $s=0,1,2,...$, and $m=0,1,2,...$.

From (4.2.3) and (4.2.6), one obtains a bounded *admissible* controller as

$$u=\mathbf{sat}_{u_{\min}}^{u_{\max}}\Big\{G_x(t)B(x)^T\sum_{i=0}^{h}\operatorname{diag}\,x^{\frac{i\ g}{2g+1}}\ K_{xi}(t)x^{\frac{i+g+1}{2g+1}}$$
$$+G_e(t)B_e^T\sum_{i=0}^{V}\operatorname{diag}\,e^{\frac{i\ b}{2b+1}}\ K_{ei}(t)e^{\frac{i+b+1}{2b+1}}+G_i(t)B_e^T\frac{1}{s}\sum_{i=0}^{s}\operatorname{diag}\,e^{\frac{i\ m}{2m+1}}\ K_{si}(t)e^{\frac{i+m+1}{2m+1}}\div\Big\} \qquad (4.2.7)$$

Here, $K_{xi}(\cdot):\ |_{\ 0}\ \ |^{c\ c}$, $K_{ei}(\cdot):\ |_{\ 0}\ \ |^{b\ b}$ and $K_{si}(\cdot):\ |_{\ 0}\ \ |^{b\ b}$ are the matrix-functions.

It is evident that assigning the integers to be zero, the well-known quadratic Lyapunov candidate results, and

$$V(t,x,e)=\tfrac{1}{2}x^TK_{x0}(t)x+\tfrac{1}{2}e^TK_{e0}(t)e+\tfrac{1}{2}e^TK_{s0}(t)e.$$

The bounded controller is found to be

$$u=\mathbf{sat}_{u_{\min}}^{u_{\max}}\Big\{G_x(t)B(x)^TK_{x0}(t)x+G_e(t)B_e^TK_{e0}(t)e+G_i(t)B_e^TK_{s0}(t)\frac{1}{s}e\div$$

Substituting (4.2.7) into (4.2.1), the total derivative of the Lyapunov candidate $V(t,x,e)$ is obtained. Solving (4.2.5), the feedback coefficients are obtained.

Example 4.2.5.

Consider a micro-electric drive actuated by a permanent-magnet DC motor with *step-down* converter, see Figure 4.2.2. Find the control algorithm.

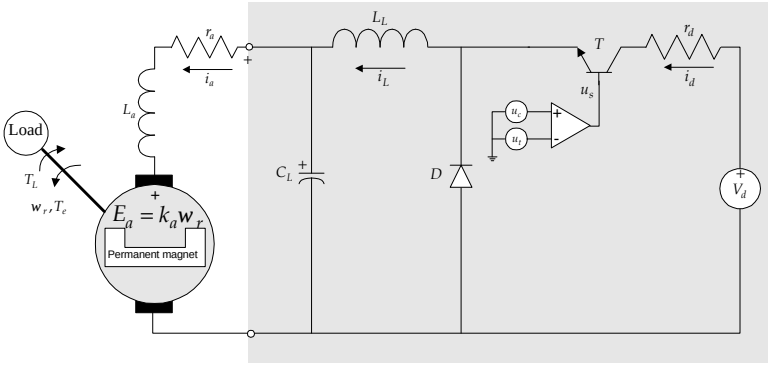


Figure 4.2.2. Permanent-magnet DC motor with *step-down* converter

Solution.

Using the Kirchhoff laws and the *averaging* concept, we have the following nonlinear state-space model with bounded control

$$\begin{matrix} \frac{du_a}{dt} \\ \frac{di_L}{dt} \\ \frac{di_a}{dt} \\ \frac{dw_r}{dt} \end{matrix} = \begin{matrix} 0 & \frac{1}{C_L} & \frac{1}{C_L} & 0 \\ \frac{1}{L_L} & 0 & 0 & 0 \\ \frac{1}{L_a} & 0 & \frac{r_a}{L_a} & \frac{k_a}{L_a} \\ 0 & 0 & \frac{k_a}{J} & \frac{B_m}{J} \end{matrix} \begin{matrix} u_a \\ i_L \\ i_a \\ w_r \end{matrix} + \begin{matrix} 0 \\ \frac{V_d}{L_L u_{t\max}} \\ 0 \\ 0 \end{matrix} \begin{matrix} 0 \\ \frac{r_d}{L_L u_{t\max}} i_L \div \\ 0 \\ \frac{1}{J} \end{matrix} \begin{matrix} \\ u_c \\ \\ \end{matrix} \begin{matrix} \\ T_L \\ \\ \end{matrix},$$

u_c [0 10] V.

A bounded control law should be synthesized.

From (4.2.6), letting $V = s = 1$ and $b = m = h = g = 0$, one finds the nonquadratic function $V(e,x)$. In particular, we apply the following Lyapunov candidate

$$V(e,x) = \tfrac{1}{2}k_{e0}e^2 + \tfrac{1}{4}k_{e1}e^4 + \tfrac{1}{2}k_{ei0}e^2 + \tfrac{1}{4}k_{ei1}e^4 + \tfrac{1}{2}[u_a \quad i_L \quad i_a \quad w_r]K_{x0} \begin{matrix} u_a \\ i_L \\ i_a \\ w_r \end{matrix},$$

where K_{x0} |4 4.

Therefore, from (4.2.7), one obtains

$$u_c = \begin{cases} 10 \text{ for } u \quad 10, \\ u \text{ for } 0 < u < 10, \\ 0 \text{ for } u \quad 0, \end{cases}$$

$$u = k_1 e + k_2 e^3 + k_3 \int edt + k_4 \int e^3 dt \quad k_5 u_a \quad k_6 i_L \quad k_7 i_a \quad k_8 w_r.$$

If the criteria, imposed on the Lyapunov pair are guaranteed, one concludes that the stability conditions are satisfied. The positive-definite nonquadratic function $V(e,x)$ was used. The feedback gains must be found by solving inequality $\frac{dV(e,x)}{dt}<0$. For example, the following inequality can be solved

$$\frac{dV(e,x)}{dt} \qquad \tfrac{1}{2}\|e\|^2 \quad \tfrac{1}{4}\|e\|^4 \quad \tfrac{1}{2}\|x\|^2 .$$

Thus, from $V(e,x)>0$ and $\frac{dV(e,x)}{dt}<0$, one concludes that stability is guaranteed.

It must be emphasized that a great number of examples in design of tracking controllers for electromechanical systems are reported in the references cited below.

Example 4.2.6.

Study the flip-chip MEMS: eight-layered lead magnesium niobate actuator (3 mm diameter, 0.25 mm thickness), actuated by a monolithic high-voltage switching regulator, $1 \quad u \quad 1$ A. A set of differential equations to model the microactuator dynamics is

$$\frac{dF_y}{dt} = \quad 9472F_y + 13740F_y u + 48593u ,$$

$$\frac{dv_y}{dt} = 947F_y \quad 94100v_y \quad 2609v_y^{1/3} \quad 2750x_y ,$$

$$\frac{dx_y}{dt} = v_y .$$

Solution.

The control authority is bounded, and hence, the control is constrained. In particular, $1 \quad u \quad 1$.

The error is the difference between the reference and microactuator position. That is,

$$e(t) = r(t) \quad y(t) ,$$

where $y(t) = x_y$ and $r(t) = r_y(t)$.

Using (4.2.6) setting the nonnegative integers to be $V = s = 1$ and $b = m = h = g = 0$, we have

$$V(e,x)=\tfrac{1}{2}k_{e0}e^2+\tfrac{1}{4}k_{e1}e^4+\tfrac{1}{2}k_{ei0}e^2+\tfrac{1}{4}k_{ei1}e^4+\tfrac{1}{2}[F_y\ v_y\ x_y]K_{xo}\begin{matrix}F_y\\ v_y\\ x_y\end{matrix}.$$

Applying the design procedure, a bounded control law is synthesized, and making use of (4.2.7), one has.

$$u=\text{sat}_{1}^{+1}\left(94827e+2614e^3+4458\ \ edt+817\ \ e^3dt\right).$$

The feedback gains were found by solving inequality

$$\frac{dV\ (e,x)}{dt}\qquad \|e\|^2\quad \|e\|^4\quad \|x\|^2\ .$$

The criteria imposed on the Lyapunov pair are satisfied. In fact,

$V(e,x)>0$ and $\dfrac{dV(e,x)}{dt}\quad 0$.

Hence, the bounded control law guarantees stability and ensures tracking. The experimental validation of stability and tracking is important. The controller is tested, and Figure 4.2.3 illustrates the transient dynamics for the position for a reference signal (desired position) $r_y(t)=4\quad 10^{\ 6}\sin 1000t$.

Figure 4.2.4 illustrates the actuator position if $r_y(t)=const=4\quad 10^{\ 6}$.

From these end-to-end transient dynamics it is evident that the desired performance has been achieved, and the output precisely follows the reference position $r_y(t)$.

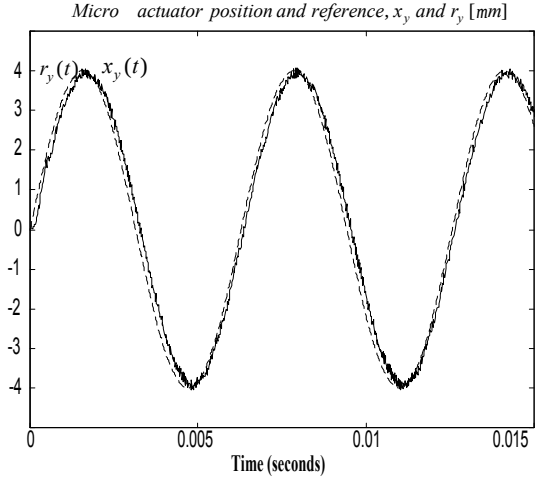


Figure 4.2.3. Transient output dynamics if $r_y(t)=4\quad 10^{\ 6}\sin 1000t$

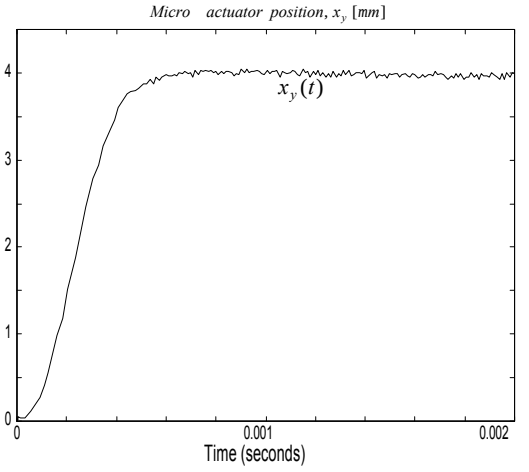


Figure 4.2.4. Actuator position, $r_y(t) = const = 4 \quad 10^{-6}$

Example 2.4.7.

Consider a flip-chip MEMS with permanent-magnet stepper motor controlled by ICs. The mathematical model in the *ab* variables, in the form of nonlinear differential equations (see section 3.6), is given as

$$\frac{di_{as}}{dt} = \frac{r_s}{L_{ss}} i_{as} + \frac{RT\psi_m}{L_{ss}} \omega_{rm} \sin(RT\theta_{rm}) + \frac{1}{L_{ss}} u_{as},$$

$$\frac{di_{bs}}{dt} = \frac{r_s}{L_{ss}} i_{bs} \quad \frac{RT\psi_m}{L_{ss}} \omega_{rm} \cos(RT\theta_{rm}) + \frac{1}{L_{ss}} u_{bs},$$

$$\frac{d\omega_{rm}}{dt} = \frac{RT\psi_m}{J} \left[i_{as} \sin(RT\theta_{rm}) + i_{bs} \cos(RT\theta_{rm}) \right] \quad \frac{B_m}{J} \omega_{rm} \quad \frac{1}{J} T_L,$$

$$\frac{d\theta_{rm}}{dt} = \omega_{rm}.$$

The two-phase micro-stepper motor parameters are:

$RT = 6$, $r_s = 60$ ohm, $\psi_m = 0.0064$ N-m/A, $L_{ss} = 0.05$ H, $B_m = 1.3 \quad 10^{-7}$ N-m-sec/rad, and the equivalent moment of inertia is $J = 1.8 \quad 10^{-8}$ kg-m^2.

The phase voltages are bounded. In particular,

$u_{\min} \quad u_{as} \quad u_{\max}$ and $u_{\min} \quad u_{bs} \quad u_{\max}$,

where $u_{\min} = -12$ V and $u_{\max} = 12$ V.

Design the tracking control algorithm.

Solution.

The nonlinear controller is given as

$$u = \begin{matrix} u_{as} \\ u_{bs} \end{matrix} = \begin{matrix} \sin(RTq_{rm}) & 0 \\ 0 & \cos(RTq_{rm}) \end{matrix}$$

$$\mathbf{sat}_{u_{\min}}^{u_{\max}} \; G_x(t)B^T \frac{\P V(t,x,e)}{\P x} + G_e(t)B_e^T \frac{\P V(t,x,e)}{\P e} + \frac{1}{s} G_i(t)B_e^T \frac{\P V(t,x,e)}{\P e} \div$$

The rotor displacement is denoted as $q_{rm}(t)$, and the output is $y(t) = q_{rm}(t)$. The tracking error is

$$e(t) = r(t) \quad y(t)$$

The Lyapunov candidate is found using (4.2.6).

Choosing a candidate Lyapunov function to be (letting $h = g = 0$ and $V = b = s = m = 1$)

$$V(e,x) = \tfrac{3}{4}K_{e0}e^{4/3} + \tfrac{1}{2}K_{e1}e^2 + \tfrac{3}{4}K_{ei0}e^{4/3} + \tfrac{1}{2}K_{ei1}e^2 + \tfrac{1}{2}\begin{bmatrix} i_{as} & i_{bs} & w_{rm} & q_{rm} \end{bmatrix} K_{x0} \begin{matrix} i_{as} \\ i_{bs} \\ w_{rm} \\ q_{rm} \end{matrix},$$

and solving

$$\frac{dV(e,x)}{dt} \qquad \|e\|^{4/3} \quad \|e\|^2 \quad \|x\|^2,$$

a bounded controller is found as

$$u_{as} = \sin(RTq_{rm})\mathrm{sat}_{12}^{+12} \; 14e + 2.9e^{1/3} + \frac{1}{s}6.1e + \frac{1}{s}4.3e^{1/3} \div,$$

$$u_{bs} = \cos(RTq_{rm})\mathrm{sat}_{12}^{+12} \; 14e + 2.9e^{1/3} + \frac{1}{s}6.1e + \frac{1}{s}4.3e^{1/3} \div$$

The sufficient conditions for robust stability are satisfied because

$$V(e,x) > 0 \text{ and } \frac{dV(e,x)}{dt} < 0.$$

Figures 4.2.5 and 4.2.6 document the dynamic if the reference (cammand) displacement was assigned 0.5 and 1 radians, respectively. From analytical and experimental results one concludes that the robust stability and tracking are guaranteed.

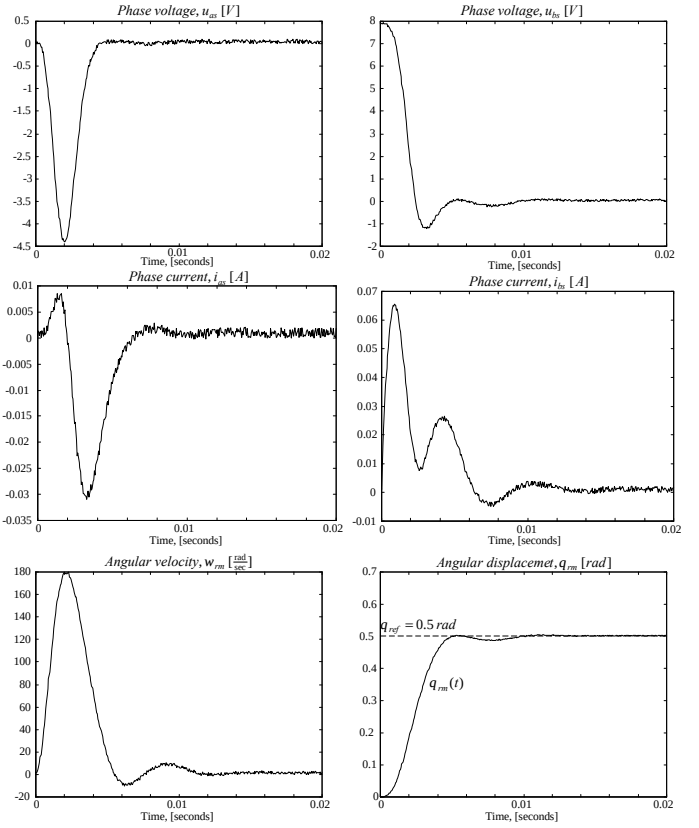


Figure 4.2.5. Transient output dynamics of microservo

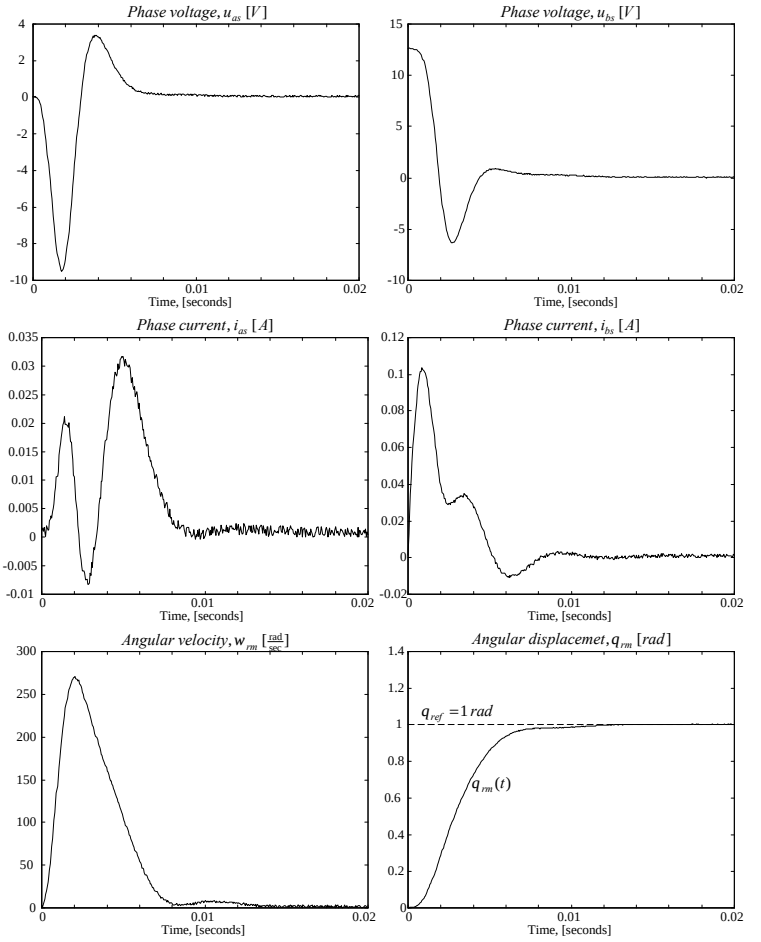


Figure 4.2.6. Transient output dynamics of microservo

References

[1] P. A. Ioannou and J. Sun, *Robust Adaptive Control*, Prentice Hall, Upper Saddle River, NJ, 1996.

[2] H. K. Khalil, *Nonlinear Systems,* Prentice Hall, Upper Saddle River, NJ, 1996.

[3] S. E. Lyshevski, *Electromechanical Systems, Electric Machines, and Applied Mechatronics*, CRC Press, FL, 1999.

[4] S. E. Lyshevski, *Control Systems Theory With Engineering Applications*, Birkh≅user, Boston, MA, 2001.

[5] M. Vidyasagar, *Nonlinear System Analysis,* Prentice Hall, Upper Saddle River, NJ, 1993.

4.3. INTRODUCTION TO INTELLIGENT CONTROL OF NANO- AND MICROELECTROMECHANICAL SYSTEMS

Hierarchical distributed closed-loop systems must be designed for large-scale multi-node NEMS and MEMS in order to perform a number of complex functions and tasks in dynamic and uncertain environments. In particular, the goal is the synthesis of control algorithms and architectures which maximize performance and efficiency minimizing system complexity through

intelligence, learning, evolution, and organization;
adaptive decision making,
coordination and autonomy of multi-node NEMS and MEMS through tasks and functions generation, organization and decomposition,
performance analysis with outcomes prediction and assessment,
real-time diagnostics, health monitoring, and estimation,
real-time adaptation and reconfiguration,
fault tolerance and robustness,
etc.

Control theory and engineering practice in the design and implementation of hierarchical hybrid (digital- and continuous-time subsystems are integrated, discrete and continuous events are augmented) real-time large-scale closed-loop systems have not matured. Synthesis of optimal controllers for elementary (single-input/single-output) single node NEMS and MEMS can be performed using conventional methods such as the Hamilton-Jacobi theory, Lyapunov's concept, maximum principle, dynamic programming, etc. However, these methods do not allow the designer to attain the desired features for complex multi-node NEMS and MEMS even though some methods (e.g., adaptive control, fuzzy logic, and neural networks) ensure performance assessment, diagnostics, adaptation, and reconfiguration. In fact, hierarchical architectures are needed to be designed and optimized to achieve intelligence, evolution, adaptive decision making, and performance analysis with outcome prediction. The design of intelligent systems can be mathematically formulated as a search problem in high-dimensional space, and the performance criteria form hypersurfaces. Efficient and robust search algorithms are used to perform optimization. Due to the complexity of large-scale systems and uncertainties, it is difficult to develop accurate analytic models, explicitly formulate performance specifications, derive regret functionals and performance indexes, design optimal architectures, synthesize hierarchical structures, as well as design control algorithms. The situation much more complex in the synthesis of robust closed-loop systems under uncertainties in dynamic environments.

Intelligence can be defined as the ability of closed-loop NEMS and MEMS achieve the desired goals (for example, maximize safety, stability, robustness, controllability, efficiency, reliability, and survivability, while minimizing failures, electromagnetic interference, and losses) in dynamic and uncertain environments through the NEMS and MEMS abilities to sense the

environment, learn and evolve, perform adaptive decision making with performance analysis and outcome prediction, and control.

Let us discuss the design of a control algorithms for *j*th level of *k*-level hierarchical NEMS and MEMS. The control law at *j*th level can be expressed as

$$u_j(t) = f\left(P, \overset{j}{\underset{i=0}{e_i(t)}}, \overset{j}{\underset{i=0}{x_i(t)}}, \overset{j}{\underset{i=0}{s_i(t)}}, p_j(t)\right),$$

where

$u_j(t)$ is the control vector (output);

f is the nonlinear map;

P is the system performance (stability, robustness, controllability, efficiency, reliability, losses, et cetera);

$\overset{j}{\underset{i=0}{e_i(t)}}$ is the error vector which represents the difference between the assigned command and events $r_i(t)$ and system outputs $y_i(t)$, and the end-to-end error vector is $e(t) = r(t) \quad y(t)$;

$\overset{j}{\underset{i=0}{x_i(t)}}$ is the state, event, and decision variable vector;

$\overset{j}{\underset{i=0}{s_i(t)}}$ is the sensed information (inputs, outputs, state and decision variables, events, disturbances, noise, parameters, et cetera) measured by *j*th and lower level sensors, and, in general, one can use $\overset{k}{\underset{i=0}{s_i(t)}}$;

$p_j(t)$ is the parameter vector (for example, time-varying parameters as well as adjustable feedback coefficients which can be changed through the decision making, learning, evolution, intelligence, control, adaptation, and reconfiguration processes).

The simplest control algorithms are the proportional-integral-derivative (PID) controllers with state/event/decision feedback extension. For example, the linear analog PID control law is given by

$$u(t) = \underset{proportional}{k_p e(t)} + \underset{integral}{k_i \frac{e(t)}{s}} + \underset{derivative}{k_d s e(t)} = k_p e(t) + k_i \int e(t)dt + k_d \frac{de(t)}{dt},$$

where k_p, k_i and k_d are the proportional, integral and derivative feedback gains.

Nonlinear PID controllers can be designed as

$$u(t)=\underbrace{\sum_{j=1}^{2N_p-1} k_{p(2j-1)}e^{2j-1}(t)}_{proportional}+\underbrace{\sum_{j=1}^{2N_i-1} k_{i(2j-1)}\frac{e(t)}{s^{2j-1}}}_{integral}+\underbrace{\sum_{j=1}^{2N_d-1} k_{d(2j-1)}\frac{d^{2j-1}e(t)}{dt^{2j-1}}}_{derivative},$$

where N_p, N_i and N_d are the positive assigned integers; $k_{p(2j-1)}, k_{i(2j-1)}$ and $k_{d(2j-1)}$ are the proportional, integral, and derivative feedback coefficients.

For $N_p=1$, $N_i=1$ and $N_d=1$, one has

$$u(t)=k_p e(t)+k_i\int e(t)dt+k_d\frac{de(t)}{dt}.$$

Assigning $N_p=2$, $N_i=2$ and $N_d=1$, the PID-type controller is expressed as

$$u(t)=k_{p1}e(t)+k_{p3}e^3(t)+k_{p5}e^5(t)+k_{i1}\frac{e(t)}{s}+k_{i3}\frac{e(t)}{s^3}+k_{i5}\frac{e(t)}{s^5}+k_{d1}\frac{de(t)}{dt}.$$

The feedback gains k_p, k_i and k_d can be time-varying.

In addition, k_p, k_i and k_d can be nonlinear functions of the state, event, and decision variables, error, disturbances, etc. For example,

$$u(t)=k_p(t)e(t)+k_i(e(t))\int e(t)dt$$
$$=\left(2+e^{-10t}\right)e(t)\int\left(3+4e^{-20t}\right)\tanh(|e(t)|)\,e(t)dt,$$

where $k_p(t)=2-e^{-10t}$ and $k_i(e(t))=\left(3+4e^{-20t}\right)\tanh(|e(t)|)$.

Assume that in large-scale multi-node NEMS/MEMS, which have thousands of nodes (NEMS/MEMS with subsystems – sensors, actuators, and ICs), one sensor and actuator were failed. These types of failures must be identified in real-time (through diagnostics and health monitoring), and closed-loop NEMS/MEMS must be reconfigurated through intelligence and adaptive decision making with performance analysis with outcome prediction and assessment. Hierarchically distributed closed-loop systems must be designed for large-scale multi-node NEMS and MEMS using hierarchical layers. For example, for three-layer configuration, the possible architecture consists of

high-level layer (intelligent augmented/coordinated control with intelligence and adaptive decision making with performance analysis and outcome prediction and assessment),

medium-level layer (intelligent coordinated or autonomous control),

low-level layer (sensor processing, data acquisition, simple feedback, e.g., single-input/single-output system for single node with many the state, event, and decision variables).

Thus, the problem can be decomposed into sub-problems performed at different layers (which can operate at different sampling rates) with synthesized layouts and decomposed tasks and functions. The system architecture must be

synthesized, and the tasks are decomposed by the high-level layer into narrow tasks/functions and fed (with or without defined adaptation, decision making, diagnostics, estimation, implementation, performance analysis, realization, as well as other details) to the medium-level layer, which further decomposes the tasks and supervises the low-level layer. This hierarchically distributed standardized control architecture is shown in Figure 4.3.1.

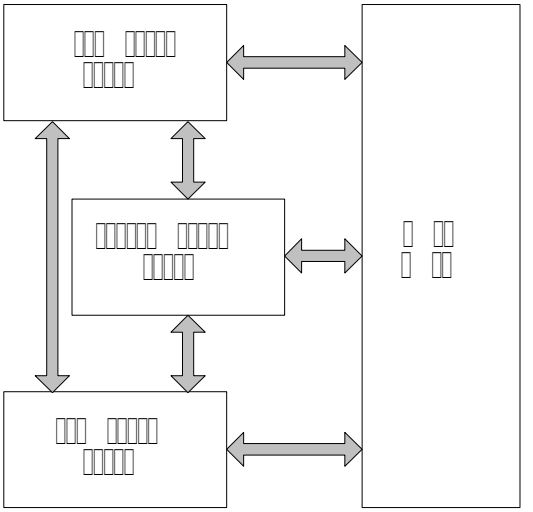


Figure 4.3.1 Three-layer hierarchically distributed architecture for large-scale multi-node NEMS and MEMS

Different operating systems, interfaces, and platforms should be supported by advanced software, and there is a critical need for novel high-performance robust software. The designer can

lay out and support hierarchical controllers in `if-then-else` execution format,

generate codes for different platforms,

add and remove layers,

set up communication and networks based upon timing requirements (write data to the shared memory buffers and read data from the buffers, protocols development, code and encode data from the buffers using different file formats),

perform diagnostics,

etc.

To perform these tasks, novel design tools are needed. At high-level, intelligence, evolution, coordination and autonomy through tasks generation/organization and decomposition, adaptive decision making with performance analysis and outcome prediction, diagnostics and estimation,

adaptation and reconfiguration, fault tolerance and robustness, as well as other functions must be performed through sensing-actuation, learning, evolution, analysis, evaluation, behavioral (dynamic and steady-state performance) optimization and adaptation, etc.

Architectures for hierarchically distributed complex closed-loop systems can be synthesized based upon the decomposition of tasks and functions. The analysis of complexity, hierarchy, data flow (sensing and actuation), and controllers design, allows the designer to synthesize architectures starting from lowest structural level and then governing and augmenting lower levels to upper levels based on physical relationships, functional correlation, order, sequence, and arrays to attain the desired performance, capability, and efficiency.

Consider the closed-loop system to displace (move) the flight control surfaces as shown in Figure 4.3.2. Ailerons, elevators, canards, flaps, rudders, stabilizers and tips are controlled by MEMS-based microactuators.

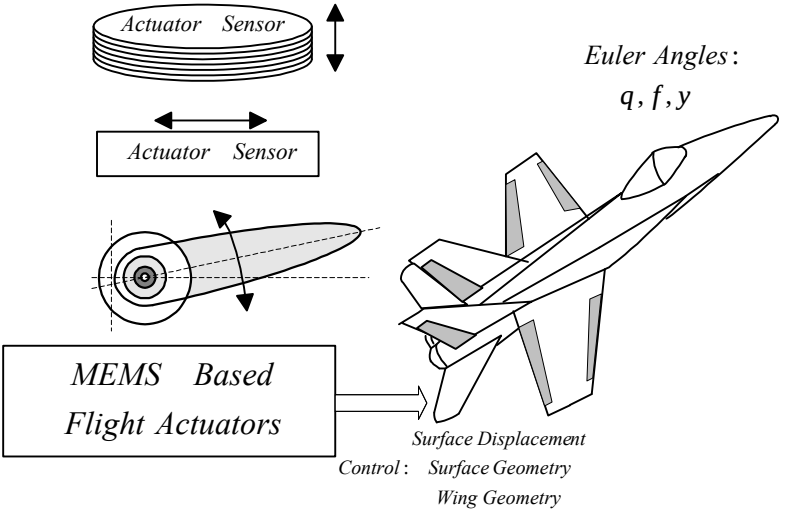


Figure 4.3.2. Aircraft with MEMS-based translational and rotational flight microactuators

Three-layer architecture of the hierarchically distributed flight management system can integrate three layers: high, medium, and low. At the low level, the single MEMS-based node (actuator/sensor) is controlled using sensing data. The control can be generated using the PID or fuzzy logic control laws (single-input/single-output system) which are based on mathematical and logical (linguistic rules to describe the system operation) reasoning. The simple tasks that can be performed by a low-level layer are:

"displace" in the clockwise or counterclockwise directions (microactuator is actuated by supplying the voltage which is controlled by microICs, and

the duty ratio of high-frequency transistors is controlled using the tracking error, decision variables, and events);

"measure the displacement",

"compare the command and actual displacements" (obtain the tracking error),

"diagnose and detect failures" (For example, positive or negative values of the duty ratio, which is bounded by ± 1, correspond to the clockwise or counterclockwise angular deflections, respectively; an increase of the duty ratio must lead to an increase of the current and electromagnetic field intensity).

Thus, we have a set of commands to attain the desired tasks and functions. The low-level layer is primarily responsible for actuation, sensing, simple analysis, diagnostics, and decision making. It is evident that the internal decision making mechanism and local diagnostics can be performed at a low level. The medium-level layer (which controls all control surfaces, e.g., left and right horizontal stabilizers, right and left leading and trailing flaps, et cetera) coordinate the actions of hundreds of MEMS-based nodes (actuators/sensors). Aircraft control is performed by a supervisory flight management system (high-level layer), which, in addition to microactuators management, integrates many other functions (e.g., mission, path, position, propulsion control, et cetera), and the task analysis is accomplished by the high-level layer. In particular, based upon the information obtained from medium- and low-level layers, the high-level layer defines tasks (through intelligence, learning, evolution, analysis, adaptation, coordination, organization, decomposition, adaptive decision making with performance analysis and outcome prediction, diagnostics, et cetera) to guarantee safety, flying and handling requirements, mission effectiveness, etc. The commands to displace the control surfaces are generated by the high-level layer based upon the overall analysis and high-level decision making. It must be emphasized that high-, medium-, and low-level layers communicate with each other, and the high-level layer possesses a key role.

Decision-making theory must be applied to develop and integrate key enabling methods, algorithms, and tools for the use in intelligent large-scale multi-node NEMS and MEMS. These intelligent systems must make optimal (robust) decision based upon the evolution strategies using specified requirements and priorities, monitoring (sensing) the external environment for entities of interest, recognizing those entities and then infer high-level attributes about those entities, etc. The closed-loop systems use the data from different sensors, feedback commands (controls) are generated and executed, and intelligent updates and evolution are performed. The feedback for sensor and control mechanisms are integrated, and particular emphasis is concentrated to gather the critical and essential data from the agents (nodes) of a greatest interest. Extensive information data must be constantly updated to guarantee a complete situation awareness, graduate evolution, and intelligence using performance analysis, outcome evaluation, prediction, and assessment. Thus,

qualitative and quantitative analysis is performed to study the overall system evolution. To perform the inferences required, to develop an assessment of the current situation, and to predict performance and outcomes, extensive knowledge and information about NEMS, MEMS, and environment are needed.

A multiple learning concept can be implemented to design high-, medium-, and low-level closed-loop systems. Agents (nodes) exhibit complex behavior which can be optimized using low-level evolutionary decision-making subsystems which use learning algorithms. Reinforcement learning can be performed based upon the prioritized objectives through upper level decision-making. The agents behavior and performance are analyzed by the high-level layer to collect and assess the evidence data. Decision trees are commonly used to provide a comprehensive set of strategies, simplify and improve (optimize) them, attain robustness and comprehensibility, and make the final decision. The low-level subsystems can perform the following functions: sensing, actuation, recognition, local diagnostics, local assessment, and local prediction with decision making.

Hierarchical distributed closed-loop systems can have different organization (architecture), and the number of layers is based upon the complexity of multi-node NEMS and MEMS to be controlled, e.g., the number of nodes their complexity. The level of control hierarchy is defined by hardware and software complexity (rate of tasks completion, rate of continuous/discrete events, bandwidth, sampling time, update rate, et cetera), as well as by the overall specifications and requirements imposed. Complex problems and tasks can be logically decomposed into simpler subproblems and subtasks which are easy to understand, support, and implement using *the state table of rules* (for each rule, the actuator or sensor action and operation are determined using *the system state, decision, event, and performance variables*). These subtasks must be performed in the defined sequence scenarios that lead to the desired operation, and the architecture is synthesized. Usually, low-level subsystem is designed for each NEMS and MEMS node (actuator/sensor) at the lower level of the hierarchy, and the layer level is defined based upon the overall objectives, analytical and numerical complexity of problems, information flow, etc. Thus, the complexity gradually arises form subsystems design (to control single node), to the synthesis of closed-loop architecture (layers).

Different intelligent concepts can be applied. For example, neural networks allow the designer to

approximate unknown functions (function approximator neural networks);

generalize control vector (control neural networks).

The backpropagation method is applied for training multilayer perceptron neural networks.

The multi-input/single-output neuron output is given by

$$u = f\left(Wv + B_1\right),$$

where u is the neuron output, u |1; f is the nonlinear function (*transfer function*); W is the weighting matrix, $W = \lfloor w_{11}, w_{12}, \ldots, w_{1k\ 1}, w_{1k} \rfloor$, W |1 k; v is the input vector (performance variables), v |k; B_1 is the bias variable.

It should be emphasized that W and B_1 are adjusted through the training (learning) mechanism.

For a single-layer neural network of z neurons, one has

$$u = f(Wv + B),$$

where the weighting matrix and bias vector are W |z k and B |z.

For multi-layer neural network of z neurons, one can find the following expression for the $(i + 1)$ network outputs

$$u_{i+1} = f_{i+1}(W_{i+1}u_i + B_{i+1}), i = 0, 1, \ldots, M \quad 2, M \quad 1,$$

where M is the number of layers in the neural network.

For example, for three-layer network, we have

$$u_3 = f_3(W_3u_2 + B_3), i = 2,\ u_2 = f_2(W_2u_1 + B_2), i = 1,$$

and $u_1 = f_1(W_1v + B_1), i = 0$.

Hence, one obtains

$$u_3 = f_3(W_3u_2 + B_3) = f_3[W_3f_2(W_2f_1(W_1v + B_1) + B_2) + B_3],$$

where the corresponding subscripts 1, 2 and 3 are used to denote the layer variables.

To approximate the unknown functions, weighting matrix W and the bias vector B must be determined, and the procedure for selecting W and B is called the network training. Many concepts are available to attain training, and the backpropagation, which is based upon the gradient descent optimization methods, are commonly used. Applying the gradient descent optimization procedure, one minimizes a mean square error performance index using the end-to-end neural network behavior. That is, using the inputs vector v and the output vector c, c |k, the quadratic performance functional is given as

$$J = \sum_{j=1}^{p} (c_j \quad u_j)^T Q(c_j \quad u_j) = \sum_{j=1}^{p} e_j^T Q e_j,$$

where $e_j = c_j \quad u_j$ is the error vector; Q |p p is the diagonal weighting matrix.

The steepest descent algorithm is applied to approximate the mean squire errors, and the learning rate and sensitivity have been widely studied for the quadratic performance indexes.

References

[1] P. J. Antsaklis and K. M. Passion (eds.), *An Introduction to Intelligent and Autonomous Control*, Kluwer Academic Press, MA, 1993.

[2] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, NJ, 1999.